INTEGRATION OF QUALITATIVE AND QUANTITATIVE DATA FOR DECISION AIDING IN PRODUCTION PLANNING

A Dissertation
Presented to
The Academic Faculty

by

Luis E. Herrera

In Partial Fulfillment
Of the Requirements of the Degree
Doctor of Philosophy in
Industrial and Systems Engineering

Georgia Institute of Technology
December 2007

INTEGRATION OF QUALITATIVE AND QUANTITATIVE DATA FOR DECISION AIDING IN PRODUCTION PLANNING

Approved by:

Dr. T. Govindaraj, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Kwok Tsui
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Leon F. McGinnis
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Sundaresan Jayaraman
School of Polymer, Textile and Fiber Engineering
Georgia Institute of Technology

Dr. Chen Zhou
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved November 5, 2007.

To the people who made me who I am:
My father, Gabriel Herrera, for his support, guidance and persistence;
My mother, Martha del Canto, for her love, her dedication, and her toughness to her family;
You my dear, for your love, your patience, your understanding, your support, and your wisdom during tough times

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. T Govindaraj, for his direction, support, ideas, encouragement, and dedication in pursuit of this research. This research would not have been possible without your help. In addition, I want to thank my committee members, Dr. Leon F. McGinnis, Dr. Chen Zhou, Dr. Kwok Tsui and Dr. Sundaresan Jayaraman for their invaluable guidance, support and constructive comments. Your participation made this research more valuable and more rewarding. I am also grateful to Dr. Gary Parker and Dr. Paul Griffin for their support during the periods I received Teaching Assistantships, these opportunities provided me with invaluable and rewarding experiences.

I would like to acknowledge the production planning staff at the Cooper Cameron Valves Corporation for facilitating the collection of data and the empirical evaluation of the software implementation. In particular, I owe a debt to Mr. Rickie Duty and Mr. Phillip Miller for their tireless efforts in facilitating and coordinating my access to the manufacturing facilities. Moreover, special thanks go to Mr. William Fontenot for his thoughtfulness, collaboration, and friendship.

Finally, I would like to thank my comrades and friends in the ISyE department and the Center for Human-Machine Systems Research. Of them, I am particularly indebted to Dr. Christine M. Mitchell for more things than I can list here. Each of you played important roles in keeping me focused through these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST SYMBOLS AND ABBREVIATIONS

DSS	Decision Support System
ERP	Enterprise Resources Planning
miniERP-GDSS	MiniERP Graph-based Decision Support System
MRP	Material Requirements Planning
MRP-II	Manufacturing Resources Planning
MTO	Make-To-Order
PICM	Production and inventory control manager
RC-MRP	Rough-cut Material Requirements Planning
SAP	*Systeme, Anwendungen und Produkte in der Datenverarbeitung* ("Systems, Applications and Products in Data Processing"). SAP is the largest business application and Enterprise Resource Planning (ERP) solution vendor and one of the biggest software providers in terms of revenues in the world.
»	Dominance relationship, e.g., “better” » “good”

SUMMARY

In this dissertation we have addressed the problem of modeling expertise in domains characterized by unquantifiable, often subjective, information, and using that model of expertise as the foundation for building computer-based decision support systems. The key feature of the expert model is to make explicit the essential characteristics of the knowledge experts use to process objective, quantitative information, for making decisions in environments rich in qualitative data. This model is then used as the basis for an "intelligent" interactive assistant that presents information appropriate for the context to operators who may not have developed the necessary expertise.

The core of the assistant is a heuristic algorithm that reflects what an expert decision maker would actually do. The algorithm incorporates a set of production rules, i.e., if-then-else rules, to define relevance conditions of quantitative data. These rules employ a *dominance principle*, i.e., a heuristic association of the relevance of quantitative data with the attributes of qualitative data, characterized as a set of ordered values. The heuristic algorithm is embedded in the assistant and is used to assist non-expert operators in locating information useful for making decisions.

The modeling methodology and the heuristic algorithm are applicable for modeling expertise in a class of decision problems characterized by large amounts of qualitative and quantitative data. The process of structuring the expert's knowledge requires empirical evidence from actual decision problems; this evidence feeds the algorithm with heuristic associations between qualitative and quantitative data. The

algorithm uses the dominance principle to decide what information to present for a particular set of conditions.

CHAPTER 1

INTRODUCTION

1.1 Background

Rapid advances in technology have led to the advent of what has been termed as the information era. Vast quantities of information are readily available through electronic and networked media on an enormous variety of subjects. This information explosion presents great challenges to decision makers who constantly find themselves in great need for efficient and effective means for accessing and utilizing the information to assess complex situations.

In addition to the complex IT systems and the growing amount of data stored in them, dynamic environments bring another level of complexity to organizations. Environments characterized by external factors (e.g., Market Trends, Technological Development, etc.) that are often unquantifiable influence strategic and tactical decisions made within the organization even when such decisions depend on data stored in the IT systems. Consequently, there is an implicit relationship and dependence between environmental factors and IT systems. In many circumstances, the characteristics of environmental factors determine the understanding and assessment of information resources.

The research described in this dissertation addresses the information needs of decision makers of modern organizations, who deal with two sources of data: IT systems and dynamic environments. Decision makers strive to make the best decisions using quantitative data (stored in IT systems) in combination with qualitative information (derived from a subjective assessment of the environment). We characterize the challenges faced by decision makers into two groups: (i) interaction with IT systems and (ii) assessment of the environment. The rapid growth and increasingly diverse types of information technology represent a challenge because of the huge amount of data stored in IT systems. These data are not necessarily readily accessible and in most cases they lack the appropriate format and content required for specific decisions. Dynamic environments bring another challenge to decision makers because of all the external factors (Market Trends, Technology), which in many situations shape the selection, understanding, and assessment of the information sources.

These challenges reflect the support needs of decision makers for using information resources and managing knowledge resources (or the lack of) associated with the assessment of the demands posed by the environment. The lack of support for these needs creates reflects in data overload problems (Fulkerson, 2000; Rondeau & Litteral, 2001).

1.1.1 Data overload and the different sources of data

Organizations are increasingly dependent on information technology (IT) systems (Drucker, 1988). IT systems include the technological resources such as hardware, software, telecommunication, and specialized human resources required to maintain records of all the organization's operations. This ubiquitous computerization of modern organizations has tremendously advanced their ability to collect, store, transform, and transmit data, producing unprecedented levels of data accumulation and access (Groover, 2002). However, the ability to extract meaning from data within individual systems or across systems has progressed much more slowly. Decision makers requiring data for specific situations find themselves bombarded with data, which often results in data overload (Woods & Patterson, 1998).

Another source of data overload is the external environment comprised of the set of relevant factors outside the boundary of an organization that are taken into consideration for the strategic and tactical decisions made in the organization (Duncan, 1972). Market conditions, consumption behavior and competency, supply chains, technological advances, political issues, to name a few, are examples of variables containing data from and about the external environment. These determine the needs and uses of data stored in the IT system for everyday operations. Data coming from external environment can be of qualitative or quantitative nature; their assessment is based on the subjective perception of the environment and depends entirely on the decision maker's expertise (domain knowledge), intuition, and situation awareness skills. In complex environments, where decisions must be made within certain time constraints, the assessment of data from the external environment can become an overwhelming task, because of the large number of factors associated with the decisions.

1.1.2 Relationship between data from the IT systems and the dynamic environment

As discussed in the previous subsections, one can recognize two sources of data involved in the decision-making process. The first set is comprised of data collected and stored by the IT system; the second set refers to data related to environmental factors. These two sets are intrinsically related during the decision-making process. The set of environmental factors provide the background or context for any decision made within the organization. On the other hand, every decision requires data stored in the IT system. The assessment of these two sets of data in conjunction with the decision making process is a complex problem because it involves various challenges: (i) the assessment of multiple types of data – qualitative and quantitative from multiple sources; (ii) the types of data exist in vast amounts and are not necessarily accessible in the required format and content; and (iii) in most cases, IT systems lack the capability to assist decision makers with information on how similar decisions have been addressed in the past. Such information, if available, could facilitate the decision making process. Organizations need to respond effectively to these challenges in order to succeed.

The relationships between the data stored in IT systems and the data emerging from the assessment of dynamic environments turns decision-making into a challenging problem with a large number of opportunities for researchers and practitioners (Bendoly, 2003). A number of questions related to this relationship have emerged. Currently, there are no clear answers to many critical questions associated with the interrelationships between the external factors and the IT System: (1) how does the rapidly changing environment affect IT design, adoption, diffusion, and assimilation in firms? (2) How should the IT system respond to the challenges posed by changes in the external factors in order to provide better support to decision makers' needs? (3) What mechanisms should be deployed to create a fit between external environment and management processes?

1.2 Problem definition

We address decision aiding in domains characterized by data overload generated by complex relationship between data stored in IT systems and data originating from the assessment of dynamic external environment. In these domains two types of data can be identified: quantitative and qualitative data. Quantitative, but unstructured and excessive data (data repositories) reflecting system status are stored and accessed through information technology (IT) systems. Qualitative (hence somewhat subjective, informed by expertise) judgment data influence the decision-making process. They can be identified and grouped into different categories of 'Environmental Factors'. Unfortunately, IT applications do not provide the means to structure and organize the effects of these factors on the decision-making process.

The types of data considered above raise two questions. The first question addresses the need to discover relevant data for a problem-solving task. The second question refers to the need of combining and integrating multiple types of information. A systematic process to structure these types of information might lead to the creation and maintenance of a memory of how decisions are made (corporate memory).

In the first question we look for alternative answers to the data overload problem, i.e., the need for decision makers to process large amounts of data. The 'data overload' phenomenon can be studied from two dimensions (perspectives): (i) origins of data, and (ii) types of data. The first dimension would consider sources of data: from external environment and from IT system sources. The second dimension considers the nature of data: (i) qualitative data and (ii) quantitative data. The assessment of the first dimension will allow us to define the nature of the data overload problem experienced by decision makers. The study of the second dimension will serve as the starting point to explore alternatives to assist decision makers' needs. A solution for the first problem will address the data overload problem that decision makers face when they deal with data originating from the external environment and from IT systems.

1.3 Proposed solution

In this dissertation we propose a way to combine and integrate, in a systematic manner, qualitative and quantitative data to aid in the decision making process. Qualitative data originates from a subjective judgment of the environment conditions. Quantitative data, stored in large databases, reflect conditions and status of system, e.g., demand forecast. We propose a decision support system that provides context-dependant help for decision makers trying to access relevant data for decisions at hand.

The core of the decision support system is a graph-based structure for qualitative and quantitative data needed for the decision-making process. The graph-based structure is comprised of two main components: (i) a model of the relationships between data repositories stored in IT systems and data originating from dynamic external environment; and (ii) a heuristic algorithm of the decision rule to gather relevant data repositories.

The first component (model) addresses the decision policies, strategies, data needs, and relationships between the environmental factors and IT data repositories. The model proposes a characterization of the different types of data repositories and categories of factors from the dynamic environment. The model then defines a relevance relationship between each data repository and each factor from the dynamic environment. The relevance relationships are created by assigning to each data repositories a pair of parameters; these parameters correspond to the importance ranking and state value of each of the external factors that make the data repository relevant. These values are represented in a two-dimensional chart.

The second component (heuristic algorithm) proposes a structured representation of the decision rules used by decision makers to identify relevant data for decisions. A special rectangular function is used to determine the conditions under which sets of data repositories are relevant under some given environmental conditions.

Together, the model and the heuristic algorithm help discover relevant data for making decisions. The graph representation itself, in the system implemented, is a visualization device; the computation of the sub-graphs comprising subsets of relevant data for different decision-making problems is based on the model and the heuristics.

The proposed solution structures the various relationships between the data based on their relevance for certain configuration of external factors. The model and the heuristic represent what expert decision makers actually do as best as possible. The implementation of the graph-based model and heuristic provides a mechanism in which data repositories are visually displayed in a graph representation. How the graph of all documents is pruned to present only those that are immediately relevant to the task at hand is based on the findings obtained during a study of actual expert decisions.

1.4 Scope

The Graph-based Model for aiding Decision Making is independent of domain of application. A concrete example, however, facilitates the description of the capabilities and components of the system. The case used to illustrate the ideas discussed in the dissertation is the Aggregate Production Planning problem in a Manufacturing domain that follows a Make-To-Order strategy.

In order to characterize the expert decision-making process for the proposed domain and problem, we conducted an ethnographic study to collect qualitative data. The qualitative research gave us valuable information on the actual decision making processes, policies, strategies, data needs, relationships between external environment and IT system. During the study we interviewed and observed senior management executives. The results of the ethnographic study were translated into an Expert Knowledge database that maps the conditions, i.e., the characteristics of the external factors (state values), under which data repositories become relevant for a decision at hand.

A computational prototype presented in proof-of-concept form implements the graph-based methodology. The prototype was designed and built to support the search for relevant data, and structure domain knowledge in production planning. The prototype includes functionality that assists the decision makers for assessing the characteristics of the environmental factors. It then maps this knowledge into the set of data repositories stored in the IT-based system to identify pieces of relevant data for making decisions (the prototype includes a module to simulate a typical SAP environment). The mapping is presented to the decision maker through an interactive graph that includes features enabling the decision maker to access relevant pieces of information.

The modeling methodology is expected to enhance decision making by allowing users to structure their domain knowledge, and to facilitate the discovery of relevant data collected and stored in the IT system data repositories. Although the computational implementation is domain dependent, we expect that the underling modeling methodology and system architecture can be easily adapted to other domains where decision makers need to deal with both qualitative and quantitative data originating from two sources: environmental factors (Market Trends, Political Issues, Economical Forces, external decision makers, etc) and from IT systems, e.g., data warehouses.

This dissertation is structured as follows. In the chapter that follows, we present a review of relevant research related to the topics and problems mentioned above. Special emphasis and interest is placed in three topics: (i) current state of research addressing the initial phase of the decision-making process, i.e., problem structuring and relevant data gathering; (ii) research addressing the theoretical and practical issues to study qualitative and quantitative data; and (iii) revision of methodologies to structure data. An ethnographic study of the decision processes of expert production planners in a manufacturing organization is presented in Chapter 3. Observations and results obtained from this study were structured to serve as the core of the graph-based modeling methodology. In Chapter 4, we present a formal representation of the modeling methodology and its two conceptual components. The first component is a model of the expert/analyst actions; the second component corresponds to a heuristic algorithm for the normative/prescriptive actions. Chapter 5 is used to describe the computational prototype and its components. The approved protocol for conducting an empirical evaluation of the computational aid and the results are described in Chapters 6 and 7, respectively. Finally, in Chapter 8, conclusions and contributions of this research are presented.

CHAPTER 2

LITERATURE REVIEW

In this chapter we provide a review and discussion of the relevant methodologies and approaches to support decision making, with special emphasis on the 'intelligence' phase (Simon, 1960). We also provide a review of the methodologies to structure data. We categorize these methodologies into two groups: (i) Methodologies to handle qualitative data, and (ii) Methodologies to handle quantitative data. We also explore existing research addressing the practical and theoretical difficulties when combining these two types of data in decision support systems. We conclude with a review of networks as a modeling construct as well as research on visualization techniques to explore networks.

2.1 Research on decision support systems

Assisting decision makers requires the use of support tools. Examples of systems that support the decision maker are expert systems (ES), executive information and support systems (EISS), and decision support systems (DSS). The nature of the management activity (Anthony, 1965), the description of the decision types (Simon, 1960), and the characteristics of the environmental forces influencing decision makers call for decision-support systems that can handle the unstructured and structured portions of the decision problem.

Decision support systems (DSS) are increasingly important tools in aiding decision makers in complex environments. The original DSS concept was most clearly defined by Gorry and Scott (Gorry & Scott Morton, 1971), who integrated Anthony's

(Anthony, 1965) categories of management activity and Simon's description of decisions (Simon, 1960). Anthony described management activities as consisting of strategic planning (executive decisions regarding overall mission and goals), tactical management (middle management guiding the organization to goals), and operational control (first line supervisors directing specific tasks). Simon described decision problems as existing on a continuum from programmed (routine, repetitive, well structured, easily solved) to non-programmed (new, novel, ill-structured, difficult to solve).

Gorry and Scott combined Anthony's management activities and Simon's description of decisions, using the terms structured, unstructured, and semi-structured, rather than programmed and non-programmed. They also used Simon's three-phase description of the decision-making process, i.e., intelligence, design, and choice. In this framework, intelligence is comprised of the search for problems, design involves the development of alternatives, and choice consists of analyzing the alternatives and choosing one for implementation. A DSS was defined as a computer-based system that dealt with a problem where at least some stage was semi-structured or unstructured. A computer system could be developed to deal with the structured portion of a DSS problem, but the judgment of the decision-maker was brought to bear on the unstructured part, hence constituting a human–machine, problem-solving system.

Gorry and Scott also argued that characteristics of both information needs and models differ in a DSS environment. The ill-defined nature of information needs in DSS situations leads to the requirement for different kinds of database systems than those for operational environments. Relational databases and flexible query languages are needed.

Similarly, the ill-structured nature of the decision process implied the need for flexible modeling environments, such as those in spreadsheet packages.

Figure 1 describes what probably came to be a more customarily used model of the decision-making process in a DSS environment. Here, the emphasis came to be on model development and problem analysis. Once the problem is recognized, it is defined in terms that facilitate the creation of models. Alternative solutions are created, and models are then developed to analyze the various alternatives. The choice is then made and implemented consistent with Simon's description. Of course, no decision process is this clear-cut in an ill-structured situation. Typically, the phases overlap and blend together, with frequent looping back to earlier stages as more is learned about the problem, as solutions fail, and so forth.

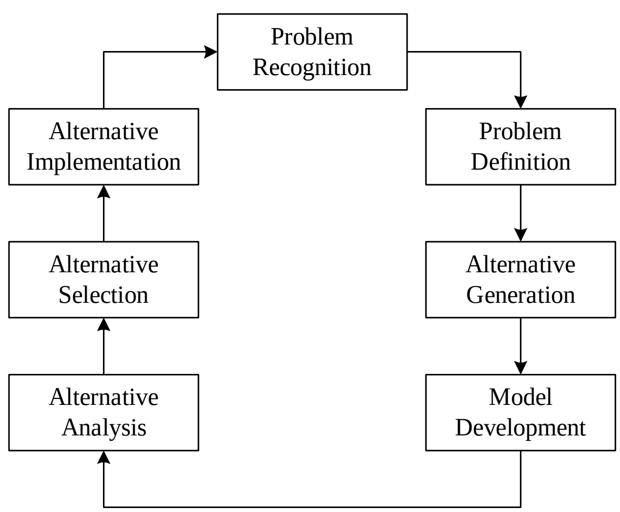


Figure 1: Customary Model of Decision Making Process

Over the last two decades, DSS research has evolved to include several additional concepts and views. Beginning in about 1985, group decision support systems (GDSS) evolved to provide brainstorming, idea evaluation, and communications facilities to support team problem solving. Executive information systems (EIS) have extended the scope of DSS from personal or small group use to the corporate level. Model management systems and knowledge-based decision support systems have used techniques from artificial intelligence and expert systems to provide smarter support for the decision-maker (Bonczek & Holsapple, 1981; Courtney & Parandise, 1993). The latter began evolving into the concept of organizational knowledge management (Paradise & Courtney, 1989) about a decade ago, and is now beginning to mature.

2.1.1 <u>The poor emphasis on the Intelligence Phase of the design of DSS</u>

The right solution to the right problem

The current state of research and practice in decision support systems provides the potential for adequate support in the information retrieval and problem analysis phases of decision-making efforts. Much less support, however, is offered for what perhaps are more critical phases of decision making: problem structuring and formulation (Huber, 1983; Lant & Hewlin, 2002; Mintzbergh & Raisinghani, 1976). Problem structuring and formulation involve postulating what the elements or variables in a problem are and how these elements fit together or interact. As pointed out by Leavitt (Leavitt, 1976), management researchers have placed far more emphasis on problem solving than on problem finding and structuring. In a related comment Schmidt states "Our business is not on answering questions, but on asking them accurately" (Schmidt, 2006). The application of DSS to semi-structured problems may lead to poor decision performance, not because incorrect solutions are derived, but because solutions to the wrong problems are sought. The discrepancy between the research focus and the research question is referred to as a type III error, one that provides the right answer for the wrong question (Murdock et al., 2007).

Semi-structured problems and types of data

Semi-structured problems cannot be defined precisely enough to use programmed solution techniques, yet have enough structure to permit effective use of computer support. These problems require human intervention, particularly during the phase of problem structuring and assessing. Pracht (Pracht, 1990; Pracht & Courtney, 1988) identified two sub-phases in the intelligence phase. “Data gathering”, the first sub-phase consists of identifying and collecting the required data. In their work, Pracht and Courtney identified two types of data: qualitative and quantitative data. “Data structuring”, the second sub-phase was defined as the stage where gathered data start acquiring some meaning for the problem at hand.

A comprehensive and generalized DSS should help the user at two levels: (a) to gather appropriate data, and (b) to structure gathered data. The first-level help should enable the user to ask the right questions within the context of semi-structured and often dynamic environment. This type of support should enable the user to identify and gather different types of data (qualitative and quantitative). The second type of help should facilitate the process of structuring the gathered data, which is of vital importance for the subsequent stages of decision support. Mintzbergh and Huber identified the task of data structuring as one of the most relevant activities for the development to DSS (Huber, 1983; Mintzbergh & Raisinghani, 1976). Yet, this activity has not received significant attention for the development of DSS. Users require support for identifying and structuring qualitative and quantitative data.

A dynamic and complex environment poses challenges for acquiring meaningful qualitative and quantitative data that are necessary for structuring the decision problem. A special challenge area of problem structuring refers to the types of data involved. Benamati classified these data into two categories: qualitative and quantitative types (Benamati et al., 1997; Ranganathan & Sethi, 2000). Different types of data call for different methodologies for handling them. In the following sections we discuss these methodologies. Special emphasis is placed on the impact of these methodologies on the two levels of support mentioned earlier.

2.2 Methodologies to handle qualitative data in decision support systems

We have divided research on qualitative data into three categories: (i) work that discusses the concept of qualitative research and its relationship with the development of decision support systems; (ii) work that uses computer-based software for handling qualitative data; (iii) finally, research that includes the morphological analysis approach.

2.2.1 Qualitative data analysis

For the purpose of this review, and in the interests of pragmatism, we will use the term 'qualitative data' as it is conventionally used within the management field: that is, to represent those techniques of data collection and analysis that rely on non-numerical data (Cassell et al., 2006). In defining "qualitative" in this way, we seek to be inclusive of a range of techniques that focus on textual data or visual images, at the same time as excluding techniques specifically involving quantification processes, which we will address in the next subsection.

Three reasons motivate our interest in studying qualitative data and their relationship with decision support systems: (i) growing volume of qualitative data overwhelming decision makers in complex environments and yet containing valuable and vital information for accurate decisions; (ii) usefulness of qualitative data elicitation as a guide to collect and assess quantitative data; and (iii) the means for mapping between qualitative data to determine selection and clustering of relevant quantitative data.

Romano (2003) described the attractiveness of qualitative data for their mapping to certain needs of quantitative data. Understanding a complex environment begins with a qualitative assessment of its characteristics. This initial assessment is followed by a selection of appropriate quantitative methods. Information obtained through qualitative techniques, “subjective, in-depth understandings of environmental factors, and the nature or structure of these environmental factors and their mapping to certain needs of quantitative data defines and distinguishes them from quantitative methods. The very richness of qualitative data has led to qualitative methods, such as open-ended surveys or self- administered questionnaires, and interviews.

Qualitative Data (QD) have a great value to offer to those seeking to extract information, knowledge and wisdom from them. QD appears valuable on the surface; however, once collected they must be effectively analyzed to yield meaningful information. Qualitative data analysis is not as straightforward as quantitative statistical analysis.

2.2.2 Qualitative data analysis software

Before computer statistical programs arrived in the 1960’s, quantitative researchers suffered from the same difficulties as qualitative research, i.e., overwhelming amount and diversity of data. This situation has changed dramatically. Since their emergence in the 1980’s, the use of computer-assisted qualitative data analysis software

(CAQDAS) has offered numerous benefits to overcome challenges associated with traditional QDA (Blismas, 2003). Surveys of QDA software and features assessments have been carried out (Barry, 1998; Blismas, 2003) and some have experimented with commercial products designed for other uses, such as text processors, databases, data indexing systems, and hypertext systems, but with little success. Many QDA tools are excellent for specific research functions in the social sciences, case studies and ethnographic studies; however, they are not designed to create a mapping between qualitative characteristics of complex environment and quantitative methods of analysis.

Despite their notable benefits, the use CAQDAS does not come without its own challenges. QDA software has a number of associated limitations and problems including, but not limited to: designer imposed biases, de-contextualization, and poor usability and inefficiency. Most CAQDAS applications were developed in social science academic programs and many started as projects to support specific doctoral student needs (Blismas, 2003). Researchers have argued, convincingly, that no one QDA program supports the entire qualitative research (QR) life cycle, rather there are categories of software designed to support specific functions within the process. Furthermore, as pointed out above, these types of applications have not been included in most of the IT systems or business intelligence applications. Dei (1993) described the initial steps of business intelligence research as highly related. We cannot ignore the need of qualitative data assessment and analysis during the design of decision support tools of users in complex environments.

2.2.3 Morphological analysis

Morphological analysis – extended by the technique of cross consistency assessment (CCA) – is a technique developed by Zwicky (1948; Zwicky, 1969) for exploring all the possible solutions to a multi-dimensional, non-quantified problem complex. This technique is similar to the approach we propose in this research, and details are provided below. As a problem-structuring and problem-solving technique, morphological analysis was designed for multi-dimensional, non-quantifiable problems where causal modeling and simulation do not function well or at all. Zwicky developed this approach to address seemingly non-reducible complexity. Using the technique of cross consistency assessment (CCA) (Ritchey, 1998), the system however does allow for reduction, not by reducing the number of variables involved, but by reducing the number of possible solutions through the elimination of the illogical solution combinations in a grid box. Details of morphological modeling and links to this research follow.

The Morphological Approach

Morphological analysis (MA) is a method for exploring all possible solutions in a complex problem space. The method was developed by Zwicky, an astrophysicist at the California Institute of Technology. Zwicky applied MA to astronomical studies and the development of jet and rocket propulsion systems. Zwicky was motivated to study the characteristics of complex problems: (i) Multi-dimensionality: A multi-dimensional problem has many interrelated aspects. For example, the problem might have to deal with financial, political and social dimensions, as a whole. (ii) Uncertainty: Aspects of complex problems are often non-quantifiable and are continuously evolving, making causal methods or simulation unsuitable. (iii) Subjectivity: There is no right or wrong solution to the problem, only better or worse solutions.

Zwicky proposed a generalized form of morphological research: "*Attention has been called to the fact that the term morphology has long been used in many fields of science to designate research on structural interrelations, for instance in anatomy, geology, botany and biology. I have proposed to generalize and systematize the concept of morphological research and include not only the study of the shapes of geometrical, geological, biological, and generally material structures, but also to study the more abstract structural interrelations among phenomena, concepts, and ideas, whatever their character might be.*" (Zwicky, 1969, p. 34).

Essentially, general morphological analysis (MA) is a method for identifying and investigating the total set of possible relationships or "configurations" contained in a given problem complex. Typology construction (Bailey, 1994; Doty & Glick, 1994) is a special case of morphological analysis. MA, however, is more generalized in form and conceptual range.

In his main work on the subject, Zwicky (1966) summarizes the five (iterative) steps of the process:

"<u>First step</u>: The problem to be solved must be very concisely formulated.

<u>Second step</u>: All of the parameters that might be of importance for the solution of the given problem must be localized and analyzed.

<u>Third step</u>: The morphological box or multidimensional matrix, which contains all of the potential solutions of the given problem, is constructed.

<u>Fourth step</u>: All the solutions contained in the morphological box are closely scrutinized and evaluated with respect to the purposes that are to be achieved.

Fifth step: The optimally suitable solutions are selected and are practically applied, provided the necessary means are available. This reduction to practice requires in general a supplemental morphological study."

The approach begins by identifying and defining the dimensions (or parameters) of the problem complex to be investigated, and assigning each of these a range of relevant "values" or conditions. A morphological box - also fittingly known as a "Zwicky box" - is constructed by setting the parameters against each other in an n-dimensional parameter space. Each cell of the parameter space contains one particular value or condition from each of the parameters, and thus marks out a particular state or configuration of the problem complex.

For example, imagine a simple problem complex, which we define as consisting of three dimensions - let us say "color", "shape" and "size". In order to conform to Figure 2, let us further define the first two dimensions as consisting of 5 discrete "values" or conditions each (e.g. color = red, green, blue, yellow, brown) and the third consisting of 3 values (size = large, medium, small). We then have 75 cells (5 x 5 x 3 = 75) in the Zwicky box, each containing 3 conditions - i.e. one from each dimension (e.g. blue, round, small). The entire 3-dimensional matrix is, in Zwicky's terms, a morphological field containing all of the (formally) possible relationships involved. Zwicky called this "complete, systematic field coverage".

Note: In our research, we referred to Zwicky's dimensions as 'Environmental factors' and to the fields' coverage or values/conditions as to "Scenario-based values".

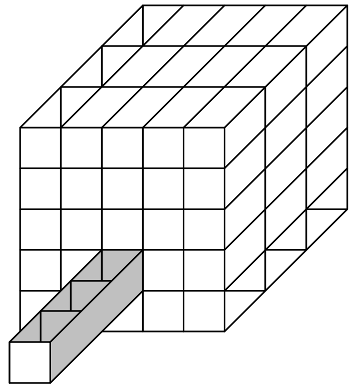


Figure 2: A 3-parameter Zwicky box containing 75 cells or “configurations”

The point is to examine all of the configurations in the field, in order to identify those that are possible, viable, practical, interesting, etc., and those that are not. In doing this, we mark out in the field what might be called a “solution space”. The “solution space” of a Zwickian morphological field consists of the subset of configurations, which satisfy some criteria.

In this research we also examine the different configurations of the fields and their coverage (environmental factors and their scenario-based values). However, our intention is to use the different configurations (solution space in Zwickian terms) as a starting point from which find useful data resources from the IT system.

The use of matrices to organize parameters, in order to uncover the multiplicity of relationships associated with a problem complex, is not new. The use of the “four-fold tables” construct, and the study of typology construction as a form of theory generation, is evidence of this fact (Bailey, 1994; Doty & Glick, 1994). However, Zwicky's highly systematic approach to this field — and his use of far more parameters than is practical in traditional typology construction — should not be underestimated. The method seeks both to be integrative and to explore the boundary conditions of complex problems. Used properly, and on the right types of problem complexes, the method is deceptively complex and rich.

The morphological approach has several advantages over less structured approaches. Zwicky calls MA "totality research" which, in an "unbiased way attempts to derive all the solutions of any given problem". It may help us to discover new relationships or configurations, which may not be so evident, or which we might have overlooked by other - less structured - methods. Importantly, it encourages the identification and investigation of boundary conditions, i.e. the limits and extremes of different contexts and factors. It also has definite advantages for scientific communication and - notably - for group work. As a process, the method demands that parameters, conditions and the issues underlying these be clearly defined. Poorly defined concepts become immediately evident when they are cross-referenced and assessed for internal consistency.

Criticisms of the MA approach

One apprehension that has been voiced against MA is that it is too structured and that this risks inhibiting free, creative thinking. For Zwicky, the whole point of morphological analysis is to get us "out of the box", to push consciousness to the limits of the conceivable and to facilitate discovery, not to obstruct it. Properly applied, general morphological analysis offers an excellent balance between freedom and (necessary) constraints. Also, computer-aided morphological analysis is a pre-eminent method for structuring and modeling what are variously called "wicked problems" and "social messes" (Ritchey, 2005a). These are complex, long-term societal and organizational planning problems that are continually evolving in a dynamic social context.

Summary

Morphological analysis, including the process of “cross-consistency assessment”, is based on the fundamental scientific method of alternating between analysis and synthesis. For this reason, it can be trusted as a useful, non-quantified method for investigating problem complexes, which cannot be treated by formal mathematical methods, causal modeling and simulation. Furthermore, both the morphological field itself, and the assessments put into the cross-consistency matrix, represent a fairly clear “audit trail”, which makes the judgmental processes inherent in MA relatively traceable, and - in a certain sense - even reproducible.

Our research has important connections with the MA Approach, e.g., we referred to Zwicky’s dimensions as ‘Environmental factors’ and to the fields’ coverage or values/conditions as to “Scenario-based values”. We also examine different configurations of the fields and their coverage (environmental factors and their scenario-based values). However, our intention is to use the different configurations (solution space in Zwickian terms) as a starting point from which to find useful data resources within the IT system.

2.3 Methodologies to handle quantitative data in decision support systems

2.3.1 Data Envelopment Analysis

Data Envelopment Analysis (DEA), as developed by Charnes et al. (1978) and extended by Banker et al. (1984) is a linear programming procedure for a frontier analysis of inputs and outputs (Andersen & Petersen, 1993). DEA is a performance measurement technique that can be used for evaluating the relative efficiency of decision-making units (DMU's) in organizations. Here a DMU is a distinct unit within an organization that has flexibility with respect to some of the decisions made, but not necessarily complete freedom with respect to these decisions. Examples of such units to which DEA has been applied are: banks, police stations, hospitals, tax offices, prisons, defense bases (army, navy, and air force), schools and university departments.

DEA had its foundations in the works of Ariat (1972), Aigner and Chu (1968), Shepard (1970), Debreu (1951), and Farrel (1957), the conceptual definitions of Koopmans (1951) and Pareto (1927), and the linear fractional transformation of Charnes and Cooper (1962) (Seiford, 1996). Early DEA-type works addressed the following issues: (i) provide LP formulations and efficient computational procedures for a variety of problems in technical efficiency including the multiple-output case; (ii) steam-electric generating plants, and (iii) aggregate census data. The first published article labeling the approach as data envelopment analysis (DEA) was Charnes, Cooper and Rhodes (Charnes et al., 1978). In 1995, DEA research had achieved a rapid growth and a widespread diffusion across multiple disciplines. For a detailed description of theoretical and practical advances see Seiford (Seiford, 1996).

Our focus in this review is on theoretical advances in regard to the inclusion of qualitative data in the DEA models. By 1990, DEA was becoming a fully developed approach. Significant advances had been made on all fronts: models, extensions, computation and practice. Theoretical refinements were extensive. Among other achievements, DEA models had been extended to include the capability to handle non-discretionary variables and / or categorical data (Banker & Morey, 1986); the ability to incorporate judgment (Dyson & Thanassoulis, 1988) and model ordinal relationships (Golany, 1988). Connections were being established with the field of decision analysis via DEA-inspired consensus ranking approaches (Cook & Kress, 1990). By 1995, DEA was recognized as a versatile and effective tool data analysis and often used as an exploratory technique (E-DEA) for visualizing data (Paradi et al., 1995).

Theoretical refinements to DEA methodology to include qualitative data

Data envelopment analysis (DEA), an effective method that is used to rank and select the best alternative from a set of alternatives, is not tailored to address qualitative criteria. The core of the DEA methodology and models searches for a performance evaluation of a series of alternatives. This evaluation is based upon a set of quantitative data. In many real world settings, however, it is essential to take into account the presence of qualitative factors when evaluating the performance of decision making units (DMU). Very often rankings are provided from best to worst relative to particular attributes. Such rank positions might better be presented in an ordinal, rather than numerical sense.

Researchers have proposed modifications to the DEA methodology to allow the inclusion of qualitative data. Zhao (2006) developed a methodology of multiple criteria data envelopment analysis (MCDEA), which can address both qualitative and quantitative criteria. MCDEA is divided into two stages for fully ranking units and each unit has multiple inputs and outputs. In the first stage, a qualitative method is applied to compare the qualitative performance of alternatives. Then, MCDEA is used to rank the alternatives by considering the relative membership degree of qualitative factors as one of the quantitative data. Cook and Zhu (2006) developed a general framework for modeling and treating qualitative data in DEA and provided a unified structure for embedding rank order data into the DEA framework. He proved a revised version of the DEA methodology for treating qualitative data. Ertay et al., (2006) proposed a modified version of the DEA methodology. He uses fuzzy linguistic variables to collect and categorize input data; analytic hierarchic process (AHP) is used to collect and organize qualitative performance data, and finally a modified DEA methodology is used to solve the decision-making problem.

2.4 Relevant methodologies for structuring data

2.4.1 Decision trees and influence diagrams

Several approaches for providing structure to external data have been used e.g., decision trees and influence diagrams. These tools offer a substantial benefit for representing decision analysis; however, they lack functionality that allows users to access critical data for specific circumstances of the domain of interest. The graphical representation requires a deeper representation of data and entities relationships.

2.4.2 Network-based representations

In modern manufacturing systems, production planning and control personnel must monitor a huge amount of high-dimensional and time-oriented data. The nature of the data sources in a manufacturing domain, and the multiple types of connections between them suggest the use of a network-based methodology to represent explicitly the rich variety of data and their interactions in a manufacturing decision-making situation. Network models facilitate the representation of the multiple entities in a decision problem, e.g., multiple forms of data, human-centered activities, personal expertise, intuition, and skills of the decision maker for any given decision at hand, etc.

In recent years, network-based modeling has emerged as a promising methodology to represent complex domains. Network-based modeling techniques have been successfully applied to represent complex problems such as socio-technological problems, the World Wide Web, geographical information systems, and medical systems.

While the network structure provides the infrastructure to represent disparate decision entities, other disciplines are required to analyze the lower-level nature of the network representation. A lower-level analysis can provide insightful discovery of hidden – or not so apparent – relationships between data. The discovery of such relationships is critical to complement the support required. Numerous disciplinary efforts have emerged to deal with this task including: (i) artificial intelligence studies oriented to support information retrieval functions; (ii) probabilistic models; (iii) statistical methods; and (iv) data mining tools.

2.4.3 Artificial Intelligence studies of network structures

There has been a great deal of research in developing methods to explore and investigate meaning of network structures. Some of these methods have been used to support information retrieval functions (Hu et al., 1999), including browsing, document clustering, spreading activation search, support for multiple search strategies, and representation of user knowledge or document content (Pracht, 1990).

2.4.4 Probabilistic models

The development of inference techniques to discover relevance properties in the elements of a network has also been an area of active research in the AI community, particularly in the context of expert systems. Two inference models based on probabilistic methods are of particular interest: Bayesian Inference Networks (Horvitz et al., 2004), and Dempster-Shafer Theory of evidence (Yager et al., 1994). These studies have tried to find some relevance in the nodes that can lead to the identification of cluster-based representations. Extensions to the Dempster-Shafer Theory of Evidence have tried to apply this approach to discover relevant information (Lalmas & Ruthven, 1998).

2.4.5 Statistical tools

Statistical models have been used for analyzing large volumes of data. In the manufacturing domain, these tools include:

Descriptive statistics: widely used to describe problems and summarize data in a useful way, either numerically or graphically. Numerical descriptors include the mean and standard deviation. Graphical summarizations include various kinds of charts and graphs.

Inference statistics: these tools have been especially helpful to model patterns in the data, e.g., discover some trends such as demand patterns, production cycles, etc. Inferences may take the form of answers to yes/no questions (hypothesis testing), estimates of numerical characteristics (estimation), prediction of future observations, descriptions of association (correlation), or modeling of relationships (regression). Other modeling techniques include ANOVA, time series, and data mining.

2.4.6 Data warehouses and On-line Analytical Processing (OLAP)

Data warehousing and on-line analytical processing (OLAP) are essential elements of decision support (Chaudhuri, 1997). Data warehousing is a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, and analyst) to make better and faster decisions. A data warehouse is a “subject-oriented, integrated, time varying, non-volatile collection of data that is used primarily in organizational decision making” (OLAPCouncil, 1997). Typically, the data warehouse is maintained separately from the organization’s operational databases. There are many reasons for doing this. The data warehouse supports on-line analytical processing (OLAP). As opposed to typical relational databases, data warehouses are targeted for decision support. Historical, summarized and consolidated data are more important than detailed, individual records. Since data warehouses contain consolidated data, perhaps from several operational databases, over potentially long periods of time, they tend to be orders of magnitude larger than operational databases.

To facilitate complex analyses and visualization, the data in a warehouse are typically modeled multi dimensionally. For example, in a sales data warehouse, time of sale, sales district, salesperson, and product might be some of the dimensions of interest. Often, these dimensions are hierarchical; time of sale may be organized as a day-month-quarter-year hierarchy, product as a product-category-industry hierarchy.

In contrast to a Data Warehouse, which is usually based on relational technology, OLAP uses a multidimensional view of aggregate data to provide quick access to strategic information for further analysis. OLAP enables analysts, managers, and executives to gain insight into data through a fast technology to create “specialized reports” and “summaries” of data. The capabilities of creating versatile reports are commonly known as “reporting features”. OLAP transforms raw data from the data warehouse so that it reflects the real dimensionality of the enterprise as understood by the user. To create specialized reports OLAP features operations such as rollup (increasing the level of aggregation), drill-down (decreasing the level of aggregation – horizontal drill-down – or increasing detail – vertical drill-down) along one or more dimension hierarchies, slice & dice (selection and projection), and pivot (re-orienting the multidimensional view of data).

We refer to data warehouses and OLAP technologies for their relevance to build the model and the proposed heuristic (see Chapter 4). During the ethnographic study (Chapter 3), we discover that decision makers experience "reports overload" as a result of the "reporting" technology included in their IT systems.

2.5 Visualization techniques to explore networks

As mentioned in the previous section, management researchers have placed far more emphasis on problem solving than on problem finding and structuring (Leavitt, 1976). Previous research efforts in decision-making processes and problem solving reveal that images and graphics are especially helpful in problem solving activities. For example, it has been shown that diagrams and graphs are useful in representing problems (Pracht & Courtney, 1988), serving as external memory aids, revealing inconsistencies in decision maker knowledge, and in leading to better understanding of novel problems and hence to improved task performance and greater decision quality.

Graphical techniques to support problem structuring have been developed in a branch of systems engineering known as structural modeling. A basic premise underlying this research is that graphical, problem-structuring tools should be incorporated into decision support systems, especially during the first phase: problem structuring and data collection. We present next some related research next.

Pracht (Pracht, 1990; 1988) proposes model visualization as an approach for business problem model development and use within a Decision Support System environment. He intended to use this approach to allow a decision maker to represent the mental images of the structure and function of a business problem model. The visualization system he proposed was based on a user-interface system that incorporated three features: (i) user-system interaction for the problem solving situation; (ii) support for discovering and modeling the structure of a complex problem; (iii) functionality to build time-based dynamic models of the complex environment.

Hu (Hu et al., 1999) conducted experiments to evaluate the impact of user-interface designs in the performance of information retrieval. Two research hypotheses drove his experiments: (i) the inclusion of graphical design concepts in an information retrieval system can help decision makers to reach "relevant" objects to the decision at hand; (ii) effective designs can alleviate the cognitive load on decision makers. Results of the experiments suggest significant implications for the design of information retrieval systems. The authors concluded that graphical interfaces were significantly better than list-based interfaces to find needed information. Another important implication is that graphical interfaces communicate information more effectively (in larger amounts and with less cognitive load).

Herrera and Komischke (2007) proposed a visualization approach and a framework to organize multidimensional data in a power generation domain. Two motives drove this work: (i) to propose a theoretical framework to organize domain knowledge to alleviate the information input overload phenomenon in process control; and (ii) to propose a set of novel concepts to design user interfaces capable of coping with the information input overload problem in process control environments.

2.6 Summary

In this dissertation research, we propose a modeling framework to support the expertise required to access relevant data for decision problems. The modeling framework brings together ideas and methodologies related to decision analysis and graph-based modeling.

A decision analysis approach is used to structure the intelligence phase of the decision making process (Simon, 1960). Decision analysis is defined as an engineering methodology in which decision maker elicits, defines, and structures the elements of the decision. We use this approach in the first phase of the decision making process: intelligence. Following this approach, we make a conceptual distinction between circumstances that are outside of control and actions we freely choose to take – 2nd Law of decision analysis. This leads us to categorize the two sets of data that intervene during the intelligence phase. Qualitative data associated to the set of environmental factors and circumstances that a decision maker needs to assess in every decision problem. The second type of data refers to data reflecting the status to systems. These data are mainly of quantitative nature.

A graph-based modeling framework is used to represent the data elements and their relationships for decision-making situations. The graph is built up in a top-down fashion. High-level elements are represented by external factors (non-controllable circumstances) to current decision. The decision maker uses his expertise and intuition (cognitive skills) to determine possible or most likely scenarios. A scenario is defined by assessing a particular configuration of external data. By defining a scenario, the network acquires a particular configuration that enables the decision maker to identify relationships useful to assess the problem at hand.

A methodology to structure decision problems characterized by large amounts of qualitative and quantitative data is presented in detail in Chapter 4. The process of structuring the expert's knowledge requires empirical evidence of actual decision process. The empirical evidence of expert's knowledge to deal with qualitative and quantitative data was collected during an ethnographic study. Details of this study are presented in the next chapter.

CHAPTER 3

DOMAIN OF APPLICATION AND ETHNOGRAPHIC STUDY

A process model for information acquisition and assessment based on studying actual experts in the field can be of great benefit for designing a decision aid capable of assisting during the intelligence phase. Such a model must be capable of representing decision-making entities and their relationships. The creation of this model is contingent upon a deeper understanding of the manner in which a decision maker thinks and reacts to problems, e.g., perceptions, cognitive responses, values, beliefs, strategies, assumptions, and data needs. An ethnographic study to develop the process model is described in this chapter.

We observed the decision-making processes of one senior management executive (expert) in a manufacturing environment. Observations occurred during five scheduled visits and covered a range of typical problems in a variety of circumstances. During each visit, the researcher observed the decision maker solving the same problem, but under different circumstances. This permitted us to capture typical variations to the problem and formulate a process model. Observations took place in the manufacturing facilities. The researcher attended meetings, observed problem-solving sessions, and interviewed the decision maker when necessary.

3.1 Domain of Application

In this section we describe in some detail the decision-making processes of one senior management executive (expert) in a manufacturing environment. From the variety of activities and decisions executed by the expert, we decided to observe only the decision-making process for the creation of the rough-cut material requirements plan (RC-MRP). Our objective was to formulate a model for the intelligence phase of the process, i.e., to capture the manner in which the expert decision maker thinks and reacts to the problem, e.g., perceptions, cognitive responses, values, beliefs, strategies, data needs, assumptions, and other subjective aspects.

A typical RC-MRP problem is solved in a period of two to three days. During this time, the decision maker defines the problem and desired objectives. He also gathers qualitative information related to the environmental factors that might have some relevance for the decision problem. This information is obtained from different personnel at various hierarchical levels of the organization, e.g., plant manager, sales, operations, logistics, materials, etc. The expert maker uses his expertise and system knowledge to collect and structure all the information received from other personnel. Finally, the decision maker retrieves and analyzes quantitative data stored in a variety of sources (mainly in the Management Information System) to create a solution for the RC-MRP problem.

The period of two to three days required to solve the problem is mainly spent in gathering qualitative data. Numerous sources participate in providing important qualitative data for the decision. The assessment of these data is an important task because it defines the characteristics of required quantitative data for solving the problem at hand.

In order to fully understand the process we scheduled five visits to the manufacturing facilities. These visits took place between the period of June 6, 2005 and April 7, 2006. Each visit lasted five days, except for the first, which required ten days to understand the dynamics of the process. The researcher attended meetings, observed the problem-solving sessions, and interviewed the expert decision maker when necessary. The decision maker was asked to act freely and to speak out his actions at all times.

Visiting the plant and observing the decision-making process in different times allowed us to observe the same process under different circumstances and conditions. This permitted the creation of a process model that could represent the most typical circumstances of the problem solving task under study. Before describing details of each visit, we present an overview of the manufacturing company, its business environment, and a characterization of the decision problem under study.

3.1.1 The Cooper Cameron Valves Corporation

The Cooper Cameron Valves (CCV) company manufactures ball valves that are used in the oil and gas industry. The CCV plant, located in Ville Platte, Louisiana, has annual sales of more than $250M and ships products to customers distributed in North America, Latin America and South America, Europe and Asia. The CCV plant produces more than 25 different types of products grouped into three categories (product lines). These groups reflect variations in the product design, type of components, and manufacturing sequence. Details of the types of products, manufacturing processes and the information management system used at the CCV are provided in Appendix A.

3.1.1.1 The dynamic characteristics of the Make-To-Order (MTO) environment

The CCV Company operates in an environment characterized by volatile demand. No seasonality could be identified in the demand of their products. In addition, the supply chain management presents serious challenges due to the long delivery times offered by raw material vendors and service provides. Moreover, the characteristics of the market indicate that it is the customers who set the delivery time of final products according to their needs. Therefore the manufacturers obtain orders for products based on their ability to produce and deliver items in required time. Variability of the required delivery time by customers poses a serious challenge to production planners in order to obtain the raw materials with appropriate anticipation to provide acceptable service (defined as the percentage of on-time deliveries).

The dynamic characteristics of the market have driven the CCV Company to choose and operate the manufacturing process following a Make-To-Order (MTO) strategy. The dynamic nature of the MTO environment makes traditional forecasting and statistical analysis tools to be of limited use for solving the forecasting problem. Moreover, the enormous amount of options available to access the data stored in the IT system and the amount of noise that comes with every report extracted result in two problems: data overload, and difficulty in discovering meaning (information) from data.

One of the challenges of the production planning in MTO manufacturing consists of making educated guesses about the future according to the environmental characteristics and likely scenarios. The decision maker requires an intelligent assessment of both the historic data and the driving external forces, to make appropriate decisions that result in the best results. A decision maker needs: (i) a structured environment (a modeling methodology) to assess the driving external forces to recognize and access the required pieces of data from the data warehouses; and (ii) an analytical methodology different from the traditional forecasting tools that permits a better assessment of future events.

The test bed application presented in this research addresses decisions that require the use of data stored in an Enterprise Resources Planning (ERP) Information Management System. The utilization of the data is dictated by the assessment of some environmental factors, e.g., economy, market conditions, customer demographics, political issues, and external decision makers' risk behavior. Data originating from environmental factors are not stored physically anywhere. The use of these data in decision making depends upon the cognitive skills of the decision makers. Naturally, the quality of the utilization of these data has an enormous variability, affecting the overall quality of the decision outcome.

Another important characteristic of these problems is that when a decision maker needs to assess a problem and access data stored in the data warehouse, these data normally are not 'clean'. Instead, the data warehouse is normally "flooded" by "dynamically-changing" data. This means that under certain circumstances, i.e., some particular characteristics of the environmental factors, some pieces of data are considered extremely "noisy" with respect to the decisions at hand; however, for other circumstances, the data could be very useful. Everything depends on the circumstances and scenarios considered by the decision maker.

3.1.2 The rough-cut material requirements plan (RC-MRP) decision

The problem of interest during the study is called the 'Rough-cut Material Requirements Planning' (RC-MRP) also known as the 'Aggregate Production Planning' (APP) decision at the Cooper Cameron Valves (CCV) plant. The output (results) of the decisions made about RC-MRP is a schedule of products that the company commits to manufacture for a planning horizon of typically six months. Table 1 displays a sample output of the RC-MRP problem. This table contains one entry for each period and each product. Each entry has a pair of values. The first value (qualitative) refers to the material supplier. The second is a quantitative value and it refers to the amount of products to be manufactured for the specific product each period.

The RC-MRP drives the purchase of raw materials, i.e., it defines the needs of raw materials for the planning horizon. It is called “rough-cut” because it actually doesn’t generate requirements for all the materials needed to build the products, but only those that qualify as “key”. Key raw materials play an important role in the production process due to the long lead time offered by raw material vendors. The importance of an accurate RC-MRP is reflected in various aspects: (i) shorter delivery times for future customers’ orders because raw material supply time constitutes an important component of the average production cycle time; (ii) it also impacts the stock levels of raw materials, when planned (and consequently manufactured) items exceed the actual demand, company incurs additional “holding costs”; and (iii) finally, when the planned quantities are smaller than the actual demand, company faces two situations, either they incur additional costs in their effort to satisfy the additional demand, or else they incur a “loss of opportunity cost”. A number of variables and factors are taken into account for the solution of the RC-MRP decision-making problem. Table 2 describes a characterization of variables involved in problem.

Table 1: Sample output for a Rough-cut Material Requirements Plan

Product size	Product family	Supplier (code) | Quantities (units)											
		Oct-01		Nov-01		Dec-01		Jan-02		Feb-02		Mar-02	
		Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty
2	02" - 06"	MSup-43573	103	MSup-43573	107	MSup-43573	111	MSup-43573	111	MSup-43573	112	MSup-43573	116
3	02" - 06"	MSup-43573	110	MSup-43573	112	MSup-43573	112	MSup-43573	117	MSup-43573	119	MSup-43573	123
4	02" - 06"	MSup-43573	110	MSup-43573	112	MSup-43573	112	MSup-43573	117	MSup-43573	119	MSup-43573	123
6	02" - 06"	MSup-43573	276	MSup-43573	282	MSup-43573	289	MSup-43573	295	MSup-43573	302	MSup-43573	308
8	08" - 12"	MSup-25789	52	MSup-25789	52	MSup-25789	52	MSup-25789	56	MSup-25789	56	MSup-25789	56
10	08" - 12"	MSup-25789	48	MSup-25789	48	MSup-25789	48	MSup-25789	52	MSup-25789	52	MSup-25789	52
12	08" - 12"	MSup-25789	48	MSup-25789	48	MSup-25789	48	MSup-25789	44	MSup-25789	44	MSup-25789	44
16	14" - larger	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53
20	14" - larger	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50
24	14" - larger	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31
30	14" - larger	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38
36	14" - larger	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31

Table 2: Characterization of the RC-MRP decision-making problem

Category	Variables
Goals	Maximize the sales group demand satisfaction
	Minimize raw material acquisition cost
	Maximize final product quality
	Minimize inventory levels and associated costs
Decisions	What raw material to purchase, i.e., which raw materials should be considered as 'key'?
	When to purchase these raw materials?, i.e., considering a delivery time and the target production goals, PICM should decide when to place orders to when should orders to raw material vendors be ?
	What quantities to purchase?
	Whom to purchase the raw materials from?
Tangible variables	Delivery time
	Quality[1]
	Price (direct cost associated with the acquisition of raw material)
Intangible variables	Raw material vendor's reliability (likelihood of variations on delivery time)
	Expedite flexibility (possibilities to shorten delivery times)
	Expedite impact on costs, i.e., extra costs associated for expediting raw material delivery
	Flexibility on raw material shipments, i.e., minimum quantities required for realizing shipments
Stakeholders	Production planning group
	Sales group
	Senior management
	Customers group
	Production shop floor group
	Inventory control group
	Raw material suppliers
	Raw material transporters
External factors	Market Trends
	Customer Stand (lead time requirements, preferences)
	Raw Materials Availability Trends
	Raw Materials Suppliers (flexibility, sales conditions)
	Corporate Policies (production goals and risk behavior)
	Competition Stand (price level and lead time)

[1] Quality has two implications; (i) one directly related to physical quality, i.e., structural composition, and (ii) the finished level of the raw material with respect to the final product, i.e., how much extra work is required to achieve the assembly level

Table 2 (continued)

Category	Variables
Data sources[2]	Demand Forecast (estimates of customers' demand needs)
	Production History (completed production orders)
	Production Backlog (current production commitments)
	Inventory Control (current inventory levels status)
	Raw Materials Supply Contracts (description of raw material suppliers analysis)
	Products Catalog (description of products and standard configuration options)
	Products' Bill of Materials (description of raw material needs for each product type)
	Products' Routing Data (description of production resources needs and manufacturing times for each product type)
	History of Raw Materials Supply Fulfilled Contracts (history of raw material supply events)
	Production Capacity: lead time, and pricing benchmark reports
Information needs[3]	Demand requirements by product, by customer, by region
	Market trends by region, by product,
	Market penetration
	Probability of receiving orders from customers of different geographical locations
	Factors that determine and increment probabilities of realizing potential customer orders: price improvement, delivery time improvement, customer-sales relationship
	Production capacities, planned vs. achieved comparisons
	Production goals (current and planned)
	Most frequent causes of production time delays
	Inventory levels by product
	Raw material supply lead time (achieved vs. planned)
	Raw material holding costs
	Customer-product opportunity costs
	Competition production-delivery-costs stands
Knowledge needs	Probabilities of realizing any particular potential order
	Negative effects and their assessed probabilities of adding a potential customer order to the aggregate production plan, e.g., failed assessment, cancellations, increments, lack of financial support
	Positive effects and their assessed probabilities of adding a potential order

[2] For the decision at hand, these are the primary sources of data that the decision maker accesses from the information management system. Decision maker refers to these as "generic documents"

[3] For the decision at hand these are the pieces of information decision maker is most interested in. Unfortunately, data obtained through the IT system (SAP R/3) do not deliver this information directly.

The person responsible for solving the RC-MRP problem is the production and inventory control manager (PICM). In the next section we describe the decision-making process followed by the PICM to solve the RC-MRP problem. In the study we observed the decision-making processes. The observations were intended to capture both the observable and non-observable actions. Observable actions include: computer transactions, data manipulations, decisions about what data to access from the IT system, and most importantly, the circumstances under which accessed data became useful for the problem at hand. We used video tape, audio recording, and occasional inquiries to the decision maker. As per actions not directly observable, these include: problem solving strategies, assumptions, preferences, etc. All these actions fall into the category of cognitive processes. Their assessment requires special techniques. We observed all the actions of the decision maker and interrupted him with questions whenever there was a need. We also asked him to speak aloud all his reasoning processes. These techniques allowed us to follow most of the decision making process. In the next chapter we describe how all this knowledge is used to formulate a model of the actions observed.

The decision making process was studied during five scheduled visits to the CCV Plant. These visits and their results are described in the subsequent sections.

3.2 The Ethnographic Study

We have two phases I and II. Phase I has two parts: A and B. In Part A, decision maker collects two types of data: (a) *qualitative* data related to environmental factors: their importance rating and their current state condition; and (b) *quantitative* data stored in databases and required to deliver a numeric solution to the problem at hand. In Part B, decision maker defines usefulness of data collected (accessed and/or generated) in Part A.

The assessment of useful (relevant) data consists of the association of each of data repository (relevant) to the characteristics (importance rating and state value) of the external factors being considered for the decision at hand.

3.2.1 Visit #1, from June 6 to June 18, 2005

3.2.1.1 Phase I: Collection of qualitative data (environmental factors)

The expert decision maker, the production and inventory control manager (PICM), initiates the process of creating the RC-MRP by gathering relevant information about the market and environment conditions. The researcher inquired the PICM about the meaning of market and environment conditions. The expert defined them as those external factors and events that have some influence on the definition of the RC-MRP. Examples of these are: customer's consumption behavior, market growth tendencies, availability of raw material, etc. The decision maker enumerated a few factors and explained that the characteristics of these factors dictate the strategies he follows to create the RC-MRP. We inquired the decision maker whether he always considered the same set of factors. The decision maker clarified that although the set of typical external factors that he used did not have a large number of elements; he typically differentiates between these factors from period to period and from problem to problem according to the internal corporate policies and his own preferences. In other words, he characterizes the set of external factors by their degree of importance and this degree is subject to change according to the corporate policies.

We inquired the decision maker regarding the types of relevant information that he gathers about each environmental factor. Decision maker stated that two types of information were relevant for each factor. The first one described above concerned the degree of importance that each factor receives. The second type was the assessment of the state value of each environmental factor. This information is received from the different personnel from inside and outside the organization. Table 3 summarizes the questions posed to the decision maker during phase 1-A of the ethnographic study.

Table 3: Decision-making process, Phase I – Part A

Researcher:	**How do you start solving the problem?**
Expert:	*"Recognize" the "market" conditions and set priorities.*
Researcher:	**What is market?**
Expert:	*Market is everything outside the company that might affect "us" For instance, "customers' behavior, "market growth tendencies", "government policies", "raw material supply availability", "competition", etc.*
Researcher:	**What is "recognize"?**
Expert:	*It means, to know what the conditions are. For example, "market tendencies", can be "good" (i.e., growing market). It also means to categorize each factor by its importance*

Observations:

1. Decision maker (DM) initiates decision-making process by assessing "environmental factors"
2. Two operations for each environmental factor are made:
   a. Categorize factors by their importance
   b. Assess the conditions of each factor
3. We observed that the assessment of each environmental factor:
   a. It lacks structure
   b. It is inconsistent

First meeting: Assessment of Market Trends and Production Goals

During the first visit we attended several meetings and conversations to understand the process of gathering the information as well as the types of information obtained. Figure 3 depicts the relationships between the decision maker and personnel from other departments.

Following his mental model of environmental factors, the decision maker decides to meet the plant manager to set the management policies and production goals. For the meeting they establish a conference phone call with the Vice President of Operations (VPO) (plant manager's boss). During the meeting they discuss several points of interest: management policies, production goals, and production constraints. VPO recalls the need to meet their budgets for the quarter. Therefore they decide to increase production levels and shipments. For that, they need to verify the last reports concerning the market tendencies. VPO invites the Vice President of Sales and Marketing (VPS&M) to the teleconference. In order to define the market growth tendencies VPS&M communicates with internal sales and district sales managers to the meeting. Together, they all agree on the current market's growing tendencies. Figure 4 depicts the information flow and actors participating in the process described above.

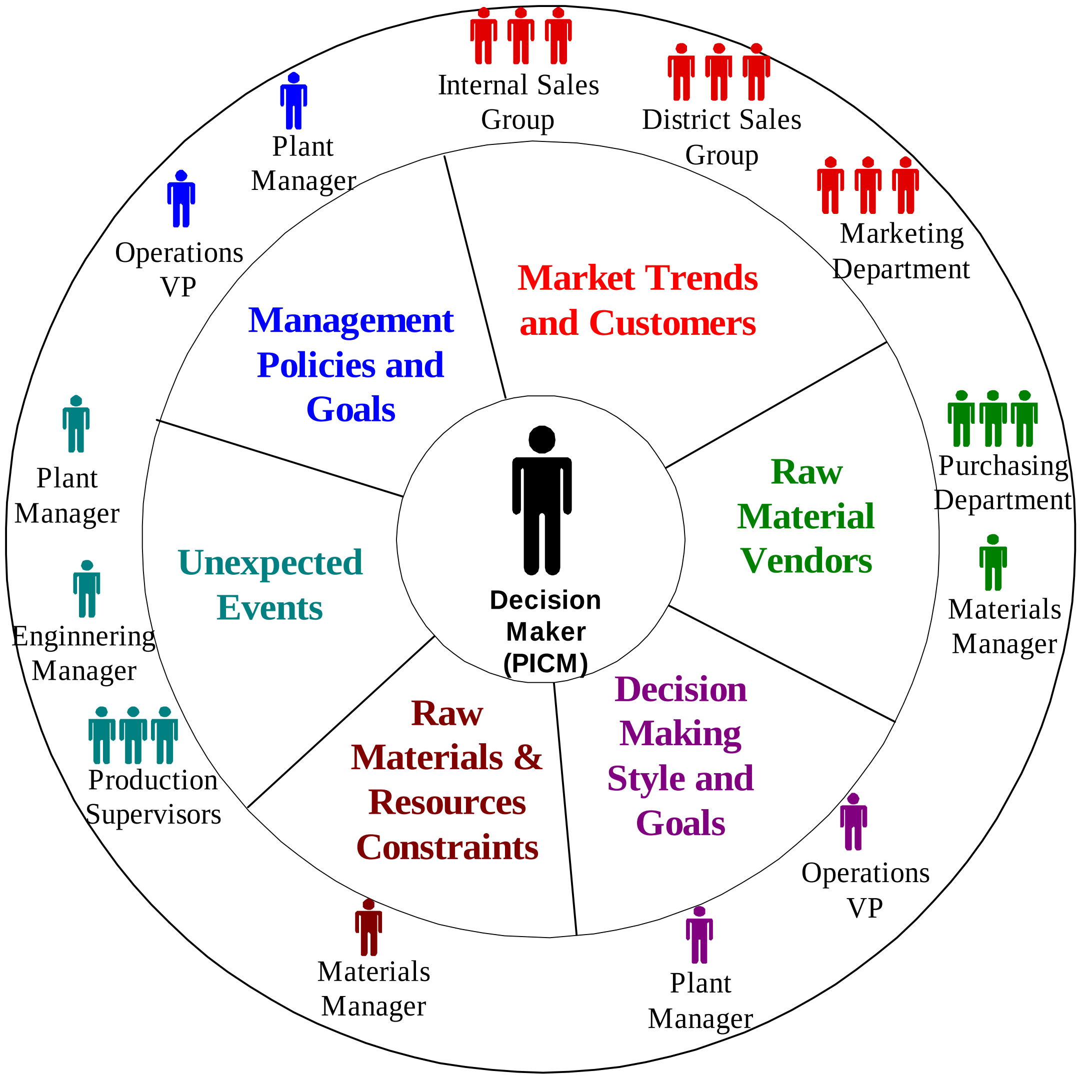


Figure 3: Information sources to assess E states

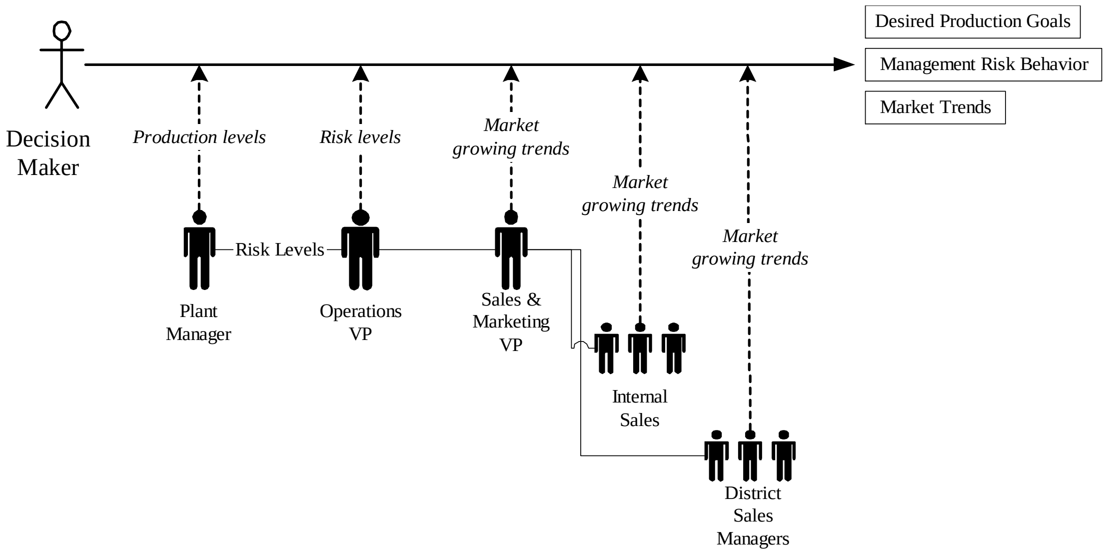


Figure 4: Information sources to assess E states for Market and Production goals

<u>Second meeting: Assessment of Material Supply</u>

Management Production Goals and Market Trends are not the only environmental factors of interest for the decision maker. The nature of the manufacturing process plays a major role in the acquisition of the raw materials and the supply chain channels. The decision maker initiates another meeting with personnel directly related to the supply of raw materials. In this meeting, the decision maker meets the materials manager and personnel from the procurement department. They discuss the most recent events concerning availability trends of raw material and raw material vendors. With respect to raw material availability, discussions focus on standard delivery conditions. Due to variations in demand, delivery times are subject to change affecting the manufacturing processes. Another aspect of raw material supply discussed in this meeting is related to expediting conditions. Material vendors review delivery conditions continuously, and in some cases the decision maker is very interested in knowing the cost and delivery time associated with improving delivery times. Figure 5 shows the information flow and the actors involved in the characterization of environmental factors related to Material Supply.

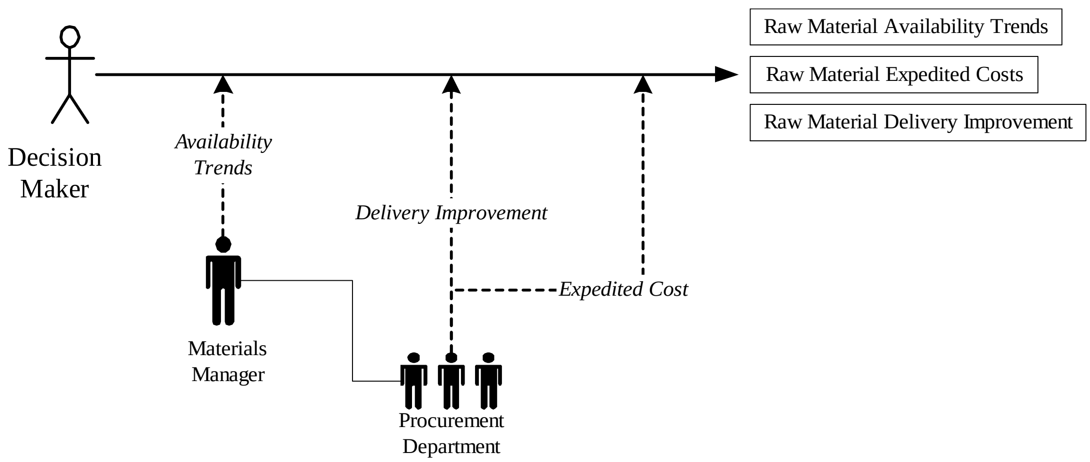


Figure 5: Information sources to assess E states for Material Supply

<u>Third meeting: Assessment of Competition and Customers state values</u>

During the third meeting, the decision maker and the plant manager meet with personnel from various sales departments. The internal sales group provides valuable input concerning the state of customer's financial position, as well as the status of customers' current orders. The decision maker needs this information to figure out the reliability level of certain customers, i.e., the likelihood of certain potential orders to become production orders. Another participant in this meeting is the Vice President of Sales and Marketing (VPS&M), who usually participates via video or telephone conference. The VPS&M provides benchmark input regarding the company's market position with regard to price and delivery time conditions. Specific details for certain customers and competitors are obtained through the district sales managers (external sales group). This group, having direct contact with customers and market conditions, is in a position to provide input concerning the required lead time conditions of customers. The decision maker uses all these inputs to characterize the state of environmental factors related to customers and competition. Figure 6 shows the information flow and actors involved in the characterization of these environmental factors.

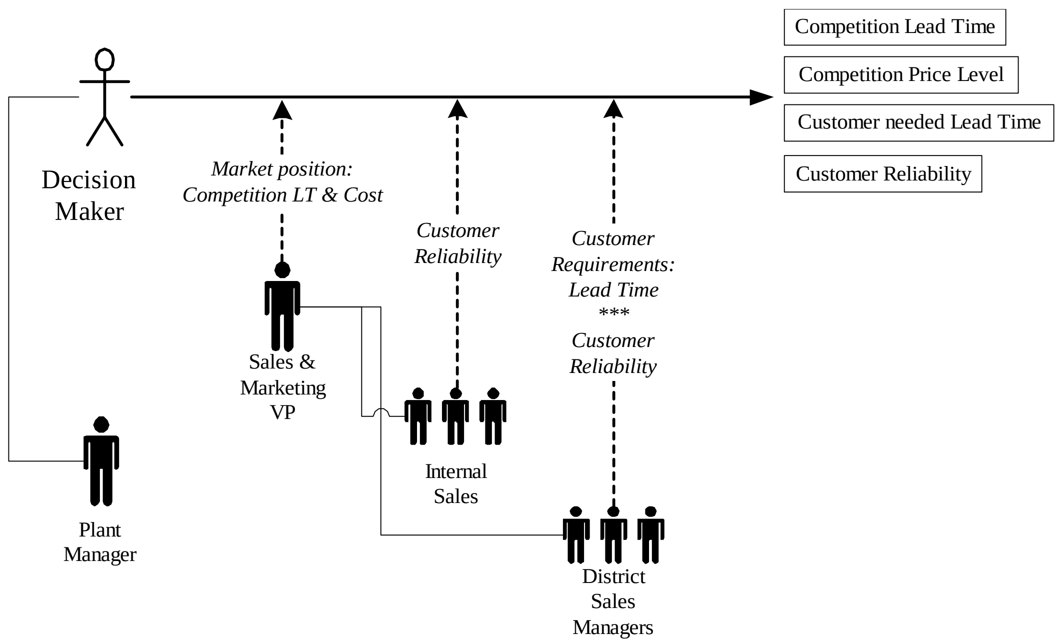


Figure 6: Information sources to assess E states for Competition and Customers

<u>Fourth meeting: Assessment of Production Conditions state values</u>

In the fourth meeting, the decision maker obtains information related to events that could potentially affect the production plan. The engineering manager provides feedback for scheduled maintenance programs of production resources. Plant manager inputs information about labor resources, i.e., scheduled labor, shifts, layoffs, etc. Production supervisors describe production conditions that could potentially affect the smoothness of production plans, e.g., potential breakdowns, shifts of bottlenecks, etc.

Finally, procurement department provides the status of raw material supply. This group estimates the likelihood of delays of raw material orders. Figure 7 depicts the information flow during this meeting.

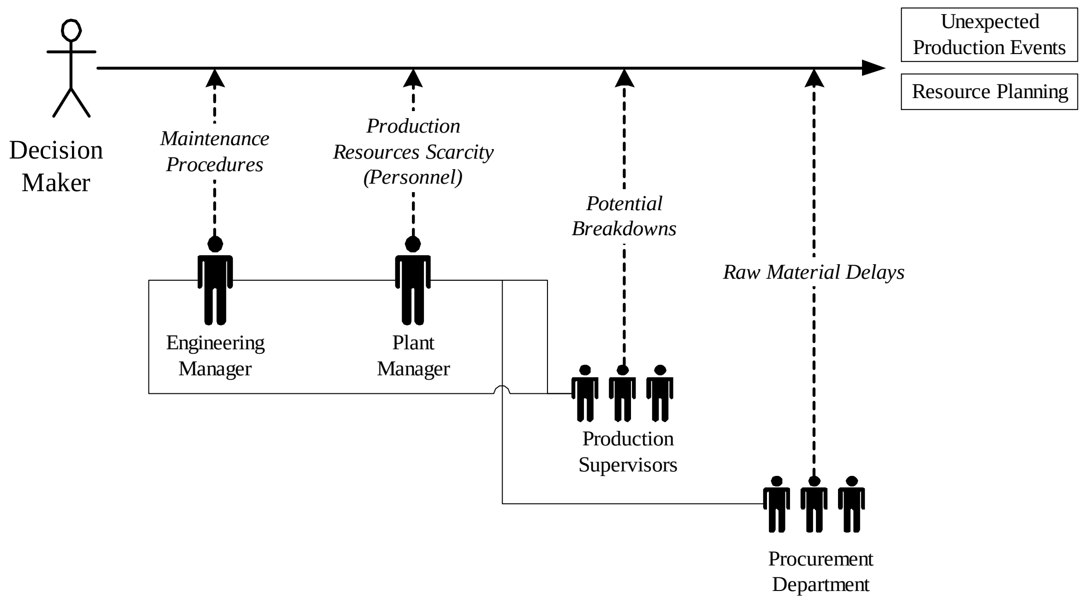


Figure 7: Information sources to assess E states for Unexpected Production Events

Assessment of environmental factors

Two types of data are assessed, quantitative data stored in databases and the IT system, and qualitative data (the state value and importance rating of environmental factors). Once the decision maker had collected all the data concerning the state of environmental factors, the researcher inquired him about the assessment of all data. Table 4 summarizes the questions posed to the decision maker during this phase of the ethnographic study.

Table 4: Decision-making process, Phase I – Part B

Researcher:	**What do you do with all these data?**
Expert:	*I use it for the data (quantitative) analysis phase*
Researcher:	**Do you structure the data (qualitative) of environmental factors?**
Expert:	*Yes, I create a summary of the relevant facts about the most important environmental factors.*
Researcher:	**How do you do it?**
Expert:	*Based on what I know about the process (my experience, knowledge, and preferences), I create a list of the most important factors to consider. For each of these factors, I ordered them according to their importance for this month. I have discussed this with my boss. Next, I assigned each factor a value based on what we learned from each factor during the interviews.*

Observations:

1. Decision maker builds a list of relevant environmental factors based on his experience and knowledge of the process. Based on the nature of the meetings, the list is consistent. Variations occur in the two grades assigned to each factor: importance and state value.
2. Decision maker judges the importance of each environmental factor.
3. Decision maker uses the input from meetings to assign a state value to each environmental factor.

Decision maker creates a summary of the results of the interviews with personnel from other departments. In that summary, the decision maker recognizes the state value of each environmental factor; this value can be either a favorable state, or a neutral state, or an adverse state. The decision maker also categorizes the list of environmental factors according to a scale of importance. The importance value is based on his knowledge of the decision-making process and it takes into account his preferences as a decision maker. Results of the assessment of environmental factors (including both importance ranking and state values) are displayed in Table 5.

Table 5: Assessment of environmental factors during visit#1: June 6, 2005

Environmental Factors	Visit #1	
	I	V
Material Delivery Improvement	3	3
Material Expedited Costs	3	2
Material Availability Trends	1	2
Competition Price Level	4	2
Competition Lead Time	3	2
Customer Needed Lead Time	4	2
Customer Reliability	2	2
Market Trends	1	2
Management Production Goals	1	3
Management Risk Behavior	1	2
Unexpected Production Events	4	3

I (importance)

***Extreme*=**	**1**
***High* =**	**2**
***Medium* =**	**3**
***Low* =**	**4**

V (State)

***Adverse*=**	**1**
***Neutral*=**	**2**
***Favorable*=**	**3**

Note: Clearly, the state-value scale is defined after the meetings with other personnel. For instance, during third meeting, the internal sales group described conditions for Customer Reliability factor as "neutral". The term "neutral" has a domain-dependent meaning. In the particular case studied, it referred to a certain level of credibility and loyalty of customers. When internal sales identify a negative tendency in this factor, they describe use the "adverse" value of the state-value scale. With respect to the importance-ranking, this value reflects the preferences and values of the company expressed through the decision maker. During the first visit, after meeting personnel from different departments and having discussed environmental conditions, the decision maker ranked all the factors according to their importance. For instance, factors such as Market Trends and Material Availability Trends received the ranking of "extreme" importance. It is important to remark (as it will be noted during the other visits) that the assessment of importance ranking is time-dependant, i.e., it changes for different times of the year. The importance ranking for all environmental factors at any point in time reflects the company’s strategic vision with respect to the environmental conditions at that point.

Phase I of the decision making process describes how the decision maker gathered and structured environmental conditions surrounding the problem-solving process. As described in Table 6, the qualitative nature of data collected from the environmental factors plays a key role during the second phase of the decision making process. The characteristics of the environmental factors drive the search and assessment of quantitative data (store in the information management system) for solving the problem at hand.

3.2.1.2 Phase II: Definition of Rules for Relevance of Quantitative Data

After all these meetings, the researcher inquired the decision maker about the usage the *qualitative* data. Decision maker explained that he uses all these *qualitative* data as guidelines to access, manipulate and analyze *quantitative* data required to create a production plan. Table 6 summarizes the interactions in the interview.

Table 6: Decision-making process, Phase II

Researcher:	**Why do you need to assess the "market"?**
Expert:	I need to do so, because it "tells me" what information (quantitative) I need to look to complete the RC-MRP
Researcher:	**How does it "tell" you?**
Expert:	It is not that simple to tell. Take for example the factor "competition". I know by experience that if competition has lower delivery times (adverse circumstances), then I need to dig into our Historic Production Report and look for opportunities to improve it. This is especially necessary if my "boss" considers the competition's lead time as an important factor.
Researcher:	**And how do you know that you need to check the historic production time?**
Expert:	I just know by experience what I need to see in most situations. There are also situations where I just have an idea of what type of information to look for. In those cases I download the entire report from 'SAP' and then play a little bit with the information until I get what I need.

Observations:

1. Decision maker (DM) recognizes a *mapping* between environmental factors and units containing quantitative data, i.e., the needs for specific types of quantitative data is determined by the characteristics of environmental factors.
2. DM maintains a mental model of the types and quality of the information he needs for each circumstance.

In order to solve the RC-MRP problem, decision maker (DM) accesses data from seven documents that are accessible from the SAP R/3 system. These documents are: the Demand Forecast Report, the Historic Production Report, the Production Backlog Report, the Raw Materials Supply Contracts Report, the Vendors Data Report, the Products' Standard Configuration, and the Production Capacity Report. Each of these reports contains data that are relevant for certain environmental factors. For instance, the Demand Forecast Report has a close relationship with Customer Needed Lead Time, Customer Reliability, Market Trends, Management Production Goals, and Management Risk Behavior. The relationship between a certain document and an environmental factor implies that the use of a document is subject to the importance and state of the environmental factor. For instance, if the importance of the environmental factor Market Trends is at least medium, then the document Demand Forecast Report is used. Similar rules apply for the rest of documents.

The usage of these documents is determined by their relationships with environmental factors. The decision maker recognizes a mapping between environmental factors (E) and quantitative data repositories (D). DM maintains a mental model of the types and quality of the information he needs for each circumstance. We can see that the mental model of what data to observe is *not very consistent* and in complex circumstances can lead the DM to *data overload*. The decision maker copes with the data overload problem using strategies such as omission and approximation.

The first step in the search of relevant documents is to look at the assessment of environmental factors. For instance, DM always starts his analysis with the Demand Forecast Report, unless the Market Trends factor was ranked with at least a “high importance” and with at least a state value of “neutral”. Table 7 summarizes the cognitive process during data manipulation of Phase II of the decision making process.

Table 7: Phase II – Data manipulation and offline analysis

Researcher:	**In the cases where you do not know in advance exactly what information to "see" how do you proceed?**
Expert:	I first look at the SAP and see if I can customize the report, if so (very unlikely), then I simply obtain the report (which in most cases comes with lots of data that I do not need). Another way to proceed (and this is the one I use most frequently) is to download the report to Excel and "play" with the data until I "see" what I need.
Researcher:	**What do you mean by "play" with data?**
Expert:	Excel has Pivot tables and pivot charts that allow me to see specific details. It is true that I need to spend some time customizing the Excel file and that I have to repeat the process every time I download the tables, but I am pretty good at that.
Researcher:	**After you complete the search of data from those pivot tables and pivot charts what happens to those files?**
Expert:	Sometimes I keep them, but with so much "stuff" going on, it's hard. Problem is that I don't keep a record of what operations I do. I sometimes find very good perspectives, but since "playing" with the pivot tables is sometimes like solving a puzzle, sometimes you get a nice solution, but it is hard to replicate.

Observations:

1. The decision maker (DM) accesses and manipulates data. In doing so, he creates multiple reports out of generic documents downloaded from IT system. The process finishes until DM finds/creates a report that contains relevant and useful information.
2. The DM uses offline analysis to find the appropriate data for each condition.
3. When the DM requires data that he cannot find directly from the information system, he downloads the entire report to a spreadsheet format and performs offline analysis.
4. Offline analysis consists of generating multiple perspectives of the data contained in the downloaded report.
5. Typical operations include: "drill down", "filter", "pivot", etc.
6. The DM recognizes two types of data obtained from the information system: level one corresponds to data as they are obtained from the system; level two corresponds to those elements of data created either through the reporting system (embedded in the SAP system) or through the "offline" analysis process by manipulating data components.
7. The DM recognizes the different levels of usefulness of each data elements. Unfortunately, he lacks a system to maintain record of each usefulness assessment.
8. The usefulness and relevance of each created report varies according to conditions (importance and state value) of external factors.

3.2.1.3 Summary of first visit

The following is a summary of the steps followed by decision maker to solve the RC-MRP combining qualitative and quantitative data and recognizing the existing mapping between the two categories of data:

a) DM recognizes the set of relevant environmental factors and their current state value. He gathers this information from his past experiences and from his conversations with personnel from other departments.
b) DM categorizes each environmental factor in two ways:
    a. Assigns a degree of importance.
    b. Recognizes the state value.
c) DM uses these assessments as guide to access both generic documents and specialized reports.
d) DM recognizes the set of available data sources. The CCV plant uses SAP R/3 as their information management system (IMS).
e) DM executes transactions to access sources of data (generic documents) from the IMS.
f) DM uses reporting technologies to create specialized reports from generic documents.
g) If reporting technology lacks functionality to create a desired specialized report, DM switched to off-line analysis. That is, he downloads data contained in generic document to a spreadsheet to create specialized reports manually and perform off-line analysis.
h) DM aborts interaction with reporting tools of the IMS.

i) During off-line analysis, DM generates multiple perspectives of downloaded document. These perspectives consist of performing several operations on data such as: rollup (increasing the level of aggregation), drill-down (decreasing the level of aggregation – horizontal drill-down – or increasing detail – vertical drill-down) along one or more dimension hierarchies, slice & dice (selection and projection), and pivot (re-orienting the multidimensional view of data) such as: horizontal and vertical drilling. These operations are easily performed using a spreadsheet environment such as the MS Excel Spreadsheets software.

j) DM differentiates between all the reports generated during the off-line analysis by assigning a level of usefulness to each data repository (specialized report) generated.

3.2.2 <u>Visits #2~#5, from August 8, 2005 to April 7, 2006</u>

Visits 2 ~ 5 were intended to capture decision maker's knowledge related to the use of documents and reports with regard to the environmental factors' importance and state values. The result of these visits is a summary of all the data repositories used for the solution of the RC-MRP problem. This summary includes the conditions for which each data repository becomes relevant (see Appendix B). These results are known as the Expert Knowledge that is used as the core for the graph-based model described in the next chapter.

3.3 Summary

In this chapter we provided a detailed description of the Cooper Cameron Valves domain of application. The description permitted a better understanding of the make-to-order operations and the type of decision involved. A detailed description of the standard manufacturing processes was provided in Appendix A.

From the variety of decisions made by the production planning staff we studied the decision processes for creating an aggregate production plan. An ethnographic study completed over five scheduled visits to the manufacturing facilities permitted us to develop a process model capable of representing typical circumstances of the problem solving task under study. The process model for information acquisition and assessment was used to build an interactive, computational tool to aid decision makers during the intelligence phase of the decision-making process.

The next chapter formalizes the process model into two conceptual components, a model of the interactions between the decision making entities and a (proposed) heuristic algorithm. Together, the model and the heuristic algorithm are intended to provide the conceptual foundation for a decision aid to structure the decision-making process for data acquisition and assessment. Chapter 5 describes the computational implementation of the model and heuristic algorithm.

CHAPTER 4

GRAPH-BASED MODEL

In this chapter we present the architecture of a graph-based model for the decision-making process identified in the ethnographic study. The core of the decision support system is a graph-based structure for qualitative and quantitative data needed for the decision-making process. The graph-based structure has two components: (i) a model of the relationships between data repositories stored in IT systems and data originating from the dynamic external environment; and (ii) a heuristic algorithm of the decision rule to select relevant data from the repositories.

The first component, the model, addresses the decision policies, strategies, data needs, and relationships between the environmental factors and IT data repositories. It characterizes the different types of data repositories and categories of factors from the dynamic environment. The model then defines a relevance relationship between each data repository and each factor from the dynamic environment. The relevance relationships are created by assigning to each data repositories a pair of parameters; these parameters correspond to the importance ranking and state value of the each external factor that make the data repository relevant. These values are represented in a two-dimensional chart.

The second component, the heuristic algorithm, proposes a structured representation of the decision rules used by decision makers to identify relevant data for decisions. A special rectangular function is used to determine the conditions under which sets of data repositories are relevant under some given environmental conditions.

Together, the model and the heuristic algorithm help discover relevant data for decision making. The graph representation itself, in the system implemented, is a means for visualization; the computation of the sub-graphs comprising subsets of relevant data for different decision-making problems is based on the model and the heuristics. Details of the model and heuristic algorithm are presented in section 4.1. Section 4.2 presents a formal (abstract) representation of the model and the heuristic.

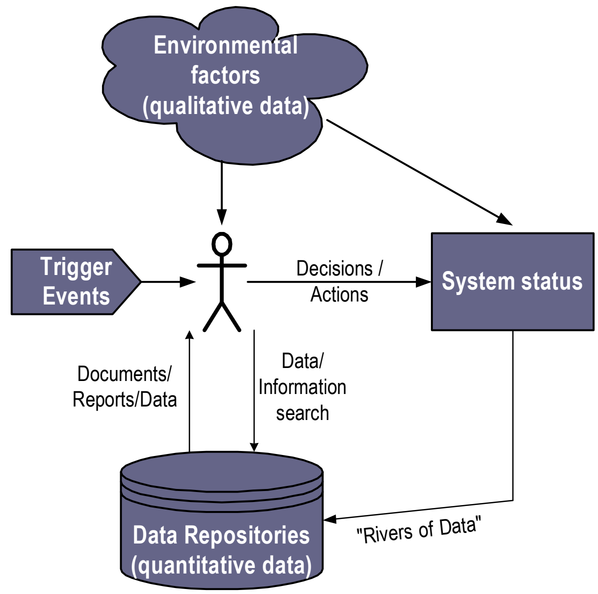
4.1 Framework of the graph-based model and heuristics algorithm

A graph representation of the decision-making entities and their relationships forms the foundation for the proposed model. For the identification of the decision-making entities, we rely on empirical observations of the phenomenon, i.e., we use observable facts from the ethnographic study. We make as few assumptions as possible to explain the phenomenon (Ockham's razor) (Wudka, 1998), or "organizing principles" in other words, i.e., infer a minimal structure for the model. Details follow.

4.1.1 Organizing principles of the model

4.1.1.1 Organizing principle #1: Identification of decision making entities

Based on the "Ockham's razor" principle, we identified the following entities: "Decision maker" (DM), "Environmental factors" (E), "Data repositories" (D) and triggering events. Relationships between these entities are defined by: decision maker's preferences, state of environmental factors for any given scenario, and conditions under which specific data elements become more useful for decision making. Figure 8 depicts the proposed framework, the decision making entities, and their relationships. In this subsection we provide details of the decision making entities. Relationships between them are described in the next subsection.



Entities	Decision maker **(DM)**	
	Environmental factors **(E)**	
	Data repositories **(D)**	
Relationships between entities	**Relationship**	**Implication**
	Connections between data repositories - Strict hierarchy	Network definition
	Decision maker - Environmental factors: (DM) - (E)	
	- Decision maker's preferences (I)	Assign relative importance rating of each E
	- State value of E for any given scenario (V)	Assess state value of each E
	Decision maker - Environmental factors/Data repositories (DM) - (D, E)	Definition of usefulness of each data repository, i.e., define environmental conditions for which data become more useful for decision making

Figure 8: Framework of the decision-making entities and their relationships

<u>Decision maker (DM)</u>

The DM is the actor that executes a series of steps to complete a problem-solving task. He uses his expertise, experience, skills and intuition to assess data from two sources: data from environmental factors (E) and data from the IT data repositories (D). Trigger events pose challenging decision making problems. His decisions affect the system status.

Environmental factors (E)

The DM needs to assess data from two sources: E and D. The determination of E depends not only on the DM's expertise and preferences, but also on the specific domain and circumstances. For instance, a production planner addressing production scheduling decisions might consider the following factors: "raw material lead time", "risk associated with changes on production orders", and "quality of raw material"; whereas another production planner in a similar domain, but under different circumstances and preferences might consider a different set of factors.

During the ethnographic study we observed that the determination of the environmental factors was an important step in the decision making process because it drives the selection of data repositories to frame a decision. For instance, a decision maker who considers Market Trends as an important environmental factor will definitely need data on Market Trends, e.g., marketing reports, sales reports, etc.

We consider two types of relationships between DM and E: (i) DM assigns a relative importance rating to each environmental factor; and (ii) decision maker assesses a value reflecting the state of current environmental factor. More details of these relationships are provided in the next subsection.

Data repositories (D)

The second source of data for the decision-making process is the set of data repositories. Data repositories can exist in different formats: text, pictures, audio, speech, and video. In this model we refer specifically to the text based data forms, and in particular to data that are accessible through an IT system. In the information age, IT systems have increased their capacity to gather and store data reflecting the organization’s operations. These capacities include reporting technologies (Chaudhuri, 1997; OLAPCouncil, 1997), i.e., tools that permit the generation of information out of data. Generated information is in the form of summarized or specialized reports.

The increased capacity for storing, distributing and generating data required a modification in the way data are organized. Information management systems, traditionally organized as local relational databases, have evolved into sophisticated distributed data warehouses, many of which hold relational databases. In most decision-making situations, analysts will need to access and assess data stored in these data warehouses. This task can become very challenging because of two reasons: firstly, data are not readily accessible in the required format for specific decisions, and secondly, the huge amount of available data and the short time required for making decisions turns the search and use of relevant data into an unfeasible task. For most problematic situations, decisions makers’ access and use of data are limited to only a few data repositories. The limited usage of data affects the quality of most decisions.

In this research we are proposing a model of data repositories and their relationships. Based on empirical evidence, the model is intended to capture two relevant facts of data: (i) the way data are stored (in a variety of relational databases); and (ii) the way summaries of data are generated through the usage of reporting technologies. The result is a graph-based representation of data repositories (nodes) and their relationships (edges). The characteristics of the node are addressed next.

There are two types of nodes in the graph representing two types of data repositories: “Generic Documents” (First–level data repositories) and “Specialized Reports” (Second–level data repositories). Figure 9 depicts these two levels. “Generic documents" are unique forms of text-based information stored and accessible through an IT system. Specialized reports are extracted data from generic reports using various "reporting technologies". These technologies are based on operations such as filtering (querying the data set), rollup (increasing the level of aggregation), drill-down (decreasing the level of aggregation – horizontal drill-down – or increasing detail – vertical drill-down) along one or more dimension hierarchies, slice & dice (selection and projection), and pivot (re-orienting the multidimensional view of data) (Chaudhuri, 1997; OLAPCouncil, 1997).

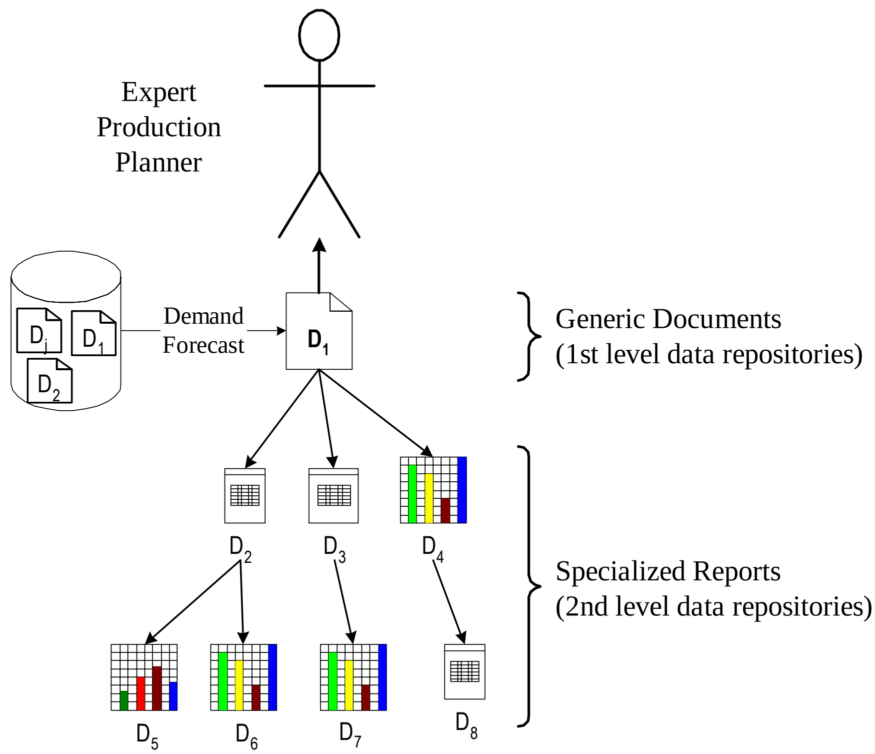


Figure 9: Graph-based representation contains two types of nodes (data repositories)

The terms 'unique' and 'generic' applied to 1st – level data repositories might need some clarification. The term 'unique' means that generic documents contain data not found in other generic documents. For instance, 'Annual Sales Report' contains data not found in another generic document such as ‘Production Routing’. The term ‘generic’ refers to data being organized in modular components containing all types of related elements. The generic document ‘Demand Forecast Report’ depicted in Table 8 contains a variety of fields reflecting important aspects of demand, e.g., “Quote_Num”, “Customer_Name”, “Book_Date”, etc. (only three projects are shown) courtesy of the Cooper Cameron Valves Corp. (Demand Forecast Data Nov. 2006).

Data repositories are divided into two types: “generic documents” and “specialized reports”. There is a parent-children relationship between these two types. A generic document can decompose into many specialized (summary) reports Details of this process will be addressed in the following subsection.

Table 8: Demand forecast data (Nov.2006) courtesy of the CCV Corp.

Field	Project 1	Projec 2	Project 3
Quote_Num	SF-CCV-21884/30	SF-CCV-98495/30	SF-CCV-65506/30
Quote_Date	2005-08-09	2005-08-16	2005-08-09
Quote_Version	3	3	3
Region	Latinamerica	Canada	Europe
District	50	8	32
Customer_Name	PETROLANE	Decatur CANADA	ALFA SAKTI
Quantity	3	4	1
Product_Code	BV-35780	BV-15818	BV-18585
Product_Size	10	10	18
Product_Line	08" - 12"	08" - 12"	14" - larger
Product_PC	ANSI 300		
Product_EndC	RFxRF	RFxRF	RFxRF
Prod Description	BV 10x10 Full	BV 10x10 Full	BV 18x18 Full
Project_Key	2	2	3
Project_Prob	75%	75%	50%
Project_Value	57 K	52 K	72K
Project_Margin	28%	27%	50%
BookDate	2005-11-07	2005-11-15	2005-11-09
Dly Rqtd (wks)	8	8	9
ShipDate	2006-01-02	2006-01-10	2006-01-11

Examples of data repositories (generic documents and specialized reports)

Consider a production planner in a manufacturing environment. Examples of data repositories might include: “Annual Sales Report”, Demand Forecast Report (see Table 8), “Production Routing Data”, etc. According to the modeling methodology, these are generic documents stored and accessible in an IT-system. These documents are independent from one another, i.e., they include all types of data related to their respective topic. For instance, a decision maker could initiate an investigation of all aspects associated with annual sales just by studying data contained in the “Annual Sales Report”. For an example of a “Specialized Report” consider the “Demand Forecast” generic document. By using certain tools such as filtering by “Project probability = 75%”, production planner could create a specialized report in which only projects with a 75% probability would be displayed. Using reporting technologies, data repositories (D) normally decompose into a finite number of sub-elements (specialized reports) showing portions of data.

4.1.1.2 Organizing principle #2: Identification of relationships between decision making entities

In the previous subsection, a set of entities involved in the decision making process were identified: “Decision Maker” (DM), “Environmental Factors” (E), “Data Repositories” (D) and triggering events. The second principle addresses the types of relationships between the entities. Three types of relationships will be studied: (i) connections between data repositories; (ii) relationships between decision makers and environmental factors; and (iii) relationships between decision makers and the coupled set of environmental factors/data repositories. Figure 8 summarizes theses relationships.

Relationships between data repositories: A discussion of the strict hierarchy of data repositories representation

Two types of data were defined in the previous subsection. We described the parent-children relationship that exists between the “generic documents” and “specialized reports”. A strict hierarchical data model is proposed to represent the parent-children relationship of data repositories (see Figure 10). In the process of searching relevant data for decisions, decision makers decompose each first-level data repository (generic document) into a finite number (N) of second-level data repositories (specialized reports). The strict hierarchy model supports the practice observed during the ethnographic study. Moreover, the strict hierarchy data model keeps a clear representation of how data are aggregated in each specialized report, i.e., it is easier to identify the process the different level of aggregation of data.

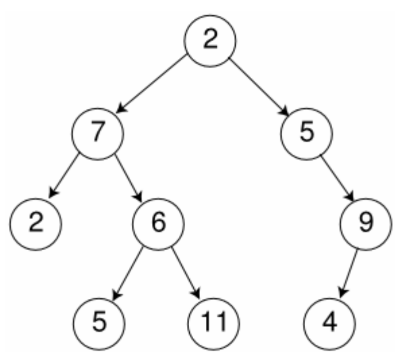


Figure 10: Strict hierarchical model of data repositories (tree structure)

Relationships between decision maker (DM) and environmental factors (E)

A decision maker assesses data from external factors (E), e.g., Market Trends, by exerting two actions on them. The first action consists of assigning a relative importance rating to each environmental factor. This rating reflects decision maker’s preferences, experience and domain knowledge for the decision at hand. This rating changes dynamically according to particular circumstances of each problem. For instance, consider a stock analyst, evaluating an investment decision. Suppose that he is assessing the relative importance of the environmental factor “technological development”. He may decide to assign a rating of ‘somehow important’ to the environmental factor.

The second action consists of assessing the state value of each environmental factor. This value reflects an evaluation of the current characteristics of the environmental factor. For instance, consider the same analyst evaluating the characteristics of the “technological development”; using his expertise and experience, he could determine that currently the “technological development is showing a growing tendency”.

The “relative importance rating” action generally assigns a qualitative value to the current environmental factor. In most cases, the scale includes values from “Not important” to “Extremely important”. On the other hand, the “state value assessment” action assigns either qualitative or quantitative values to the factor. For instance, the environmental factor Market Trends could take qualitative values ranging from “growing tendencies” to “decreasing tendencies”; whereas another environmental factor such as the “value of NYSE index”, could take numeric values such as 130 points.

The combination of qualitative and quantitative data imposes certain requirements on the modeling process. The model and heuristic algorithm proposed in this research addresses these requirements.

Relationships between decision maker (DM) and the coupled set of environmental factors/data repositories

In this relationship, decision maker provides a usefulness measure of data repositories with respect to each environmental factor. In other words, with this relationship decision maker specifies the circumstances under which each data repository becomes relevant. Additionally, this relationship provides the mechanism to keep a record of how decisions are made. In order to understand better this relationship we provide an example.

Consider a decision maker in a manufacturing environment facing a production planning decision. Suppose that one of the environmental factors that he is considering in his analysis is “Delivery times required by customers”. Assume that he is currently using four data repositories: one generic document (D_1) and three specialized reports ($R_{1,1}$, $R_{1,2}$, and $R_{1,3}$). Figure 11 depicts these entities. The ‘usefulness’ assessment of each data repository consists of defining the circumstances (relative importance and state value) of the environmental factor E_1 that will make each data repository become relevant. For instance, suppose that decision maker assigns a usefulness rating of “negotiable” to data repository $R_{1,1}$. This means that $R_{1,1}$ could become relevant, i.e. useful for current decision when environmental factor E_1 receives a state-value assessment of “negotiable”.

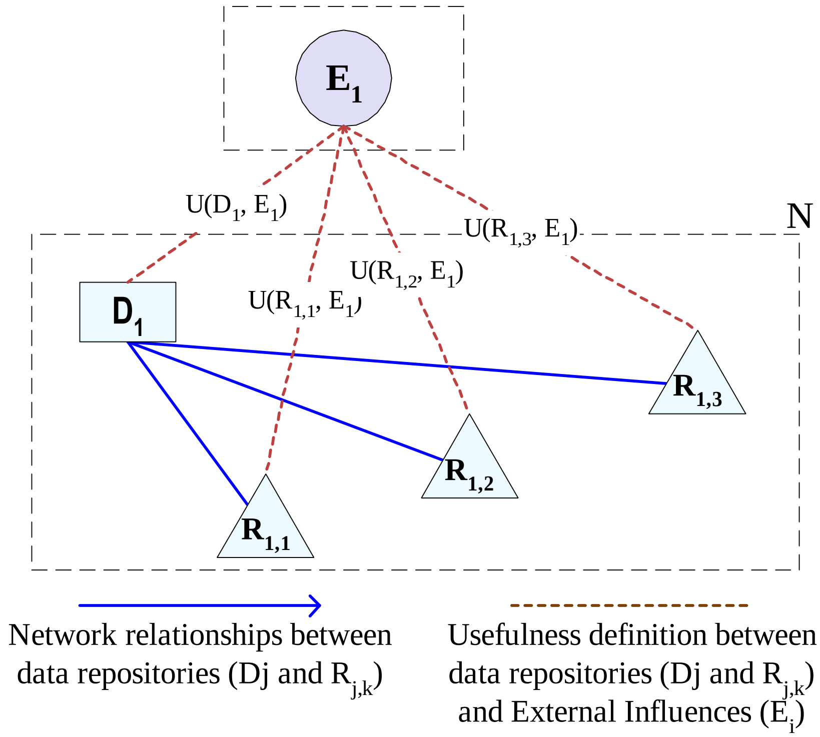


Figure 11: Usefulness is defined for each 2-tuple (D_i, E_j)

4.1.2 Dominance relationships

The objective of the heuristic algorithm is to determine when the data repositories become relevant for the decision at hand. It does this by considering State Assessment actions (importance rating and scenario-based value assessment) of environmental factors E_j. It also considers usefulness ratings assigned to the 2-tuple set of data repositories and external factor (D_i, E_j). A special rectangular function is used to determine the conditions under which sets of data repositories are relevant under some given environmental conditions.

There exist three cases in which we compare the usefulness value with the state value of the environmental factors. For completeness we will illustrate the three cases. However, as it was determined during the ethnographic study, the applicable case for the conditions of the domain of application is the second case "Positive dominance" (subsection 4.1.2.2).

Consider three categories of tasks ranked in a scale of difficulty as ['hard' (very difficult), 'medium' (normal difficulty), 'easy' (low difficulty)] and three types of tools ['A', 'B', 'C']. Consider that tool 'A' is suitable to execute task marked as 'hard', that tool 'B' is suitable to execute task marked as 'medium', and that tool 'C' is suitable to execute task marked as 'easy'. Then we ask: "for what types of tasks can each tool be used?" To answer this question, we propose three cases:

4.1.2.1 Case 1: Strict equivalence "Use tools only for the tasks for which they are suitable".

We represent this relationship in Figure 12.

Task category	Tools		
	A	**B**	**C**
Difficult	Opt	✕	✕
Medium	✕	Opt	✕
Easy	✕	✕	Opt

✕ : Not applicable

Opt : Optimal utilization

Figure 12: Strict equivalence – neutral dominance relationship

The use of a strict equivalence comparison between a usefulness value and a state value defines the circumstances under which a data repository becomes relevant. Let us apply the strict equivalence comparison to the example stated above and see the effects on the relevance of data repository.

Following our example, consider that the environmental factor E_1 corresponds to "Delivery times required by customers", and that $R_{1,2}$ corresponds to "Report of expeditable raw materials including discount levels". Suppose that the range of state values that E1 can take is [long, medium, short], notice that these values are ordered from more favorable to more adverse conditions [favorable, medium, adverse]. Suppose that from previous decisions, decision maker assigned a usefulness measure to $R_{1,2}$ with regard the E_1 as 'Short'. Finally consider that while assessing the characteristics of environmental factors, decision maker assesses the state value of E_1 as 'Short', i.e., using his experience and knowledge of the current environment, he knows that customers are currently requiring 'short delivery times.

$$\text{Usefulness } (R_{1,2} \mid E_1) = \text{'Short'}$$

$$\text{State value assessment } (E_1) = \text{'Short'}$$

According to the definition of the strict equivalence comparison, the specialized report $R_{1,2}$ becomes relevant because of the equivalence between the two values:

$$\text{Usefulness } (R_{1,2} \mid E_1) = \text{State value assessment } (E_1)$$

If values are not equal, then relevance requirements are not met and the specialized report $R_{1,2}$ is left out of the relevant network.

4.1.2.2 Case 2: Positive dominance

This is the applicable case for the conditions of the domain of application. Refer to section 4.2 for more details. The positive dominance states that if a tool is useful (relevant) for certain tasks it will also be useful under *less* demanding tasks. Figure 13 depicts this relationship.

Task category	Tools		
	A	**B**	**C**
Difficult	Opt	✗	✗
Medium	✓	Opt	✗
Easy	✓	✓	Opt

✗ : Not applicable

✓ : Utilization is ok (dominance)

Opt : Optimal utilization

Figure 13: Positive dominance relationship

The use of a positive dominance comparison between a usefulness value and a state value defines a set of values (as opposed to only one value for the strict equivalence case) or circumstances under which a data repository becomes relevant. Let us apply the positive dominance comparison to the example stated above and see the effects on the relevance of data repository.

Following the same example and the same assessments of usefulness and state values:

$$\text{Usefulness } (R_{1,2} \,|\, E_1) = \text{‘Short’}$$

$$\text{State value assessment (E1)} = \text{‘Short’}$$

According to the definition of positive dominance comparison, the assessment of usefulness:

$$\text{Usefulness } (R_{1,2} \,|\, E_1) = \text{‘Short’}$$

Defines a range of values for which specialized report $R_{1,2}$ becomes relevant. The range of values include: [long, medium, short], this is so because, if $R_{1,2}$ is relevant (useful) for 'Short' conditions it will also become relevant (useful) for less demanding conditions which include 'medium' and 'long' delivery times.

4.1.2.3 Case 3: Negative dominance

If a tool is useful (relevant) for certain tasks it will also be useful under *more* demanding tasks, but not for less demanding tasks. Figure 14 depicts this relationship.

Task category	Tools		
	A	**B**	**C**
Difficult	Opt	✓	✓
Medium	✗	Opt	✓
Easy	✗	✗	Opt

✗ : Not applicable

✓ : Utilization is ok (dominance)

Opt : Optimal utilization

Figure 14: Negative dominance relationship

The use of a negative dominance comparison between a usefulness value and a state value defines a set of values (not just one as in the strict equivalence case) circumstances under which a data repository becomes relevant. Let us apply the negative dominance comparison to the example stated above and see the effects on the relevance of data repository.

Following the same example and the same assessments of usefulness and state values:

$$\text{Usefulness } (R_{1,2} \mid E_1) = \text{‘Short’}$$

$$\text{State value assessment } (E_1) = \text{‘Short’}$$

According to the definition of negative dominance comparison, the assessment of usefulness:

$$\text{Usefulness } (R_{1,2} \mid E_1) = \text{‘Short’}$$

Defines a range of values for which specialized report $R_{1,2}$ becomes relevant. The range of values includes only the ‘short’ delivery time value because ‘short’ is the *most* demanding condition.

4.2 Formal representation of the model and the heuristic algorithm

In this section we present a formal representation of the model and the heuristic algorithm. Our intention is to propose a generalizable notation that encapsulates the model entities and their relationships as well as the decision rules contained in the heuristic algorithm.

4.2.1 Model formulation

4.2.1.1 Assumptions

Consider the situation where an *organization* is facing a *decision-making problem* (P) with the following characteristics:

- Decisions required to solve P require access to multiple types of data repositories.
- Decisions made in the organization require an accurate recognition and an intelligent judgment of environmental factors.

We assume P is solved by a *decision maker* (DM) who possesses the following characteristics:

- DM possesses a level of domain knowledge and experience that allows him to recognize and assess environmental factors to the problem at hand (P).
- DM has access to all data repositories to solve P.
- DM possesses the required expertise that allows him to evaluate (or assess) the usefulness of each data repository using the state value of the environmental factors as a reference.

Consider the set of *environmental factors* (E) influencing the organization's operations and actors:

- Information reflecting the state value of E at any point in time is accessible by the decision maker.
- Elements of E can receive a qualitative rating reflecting their state value.
- It is possible to assess external factors according to their importance to P.

We assume the existence of an *information technology system* (IT) with the following characteristics:

- IT is organized as a hierarchical[4] structure of data repositories (D).
- Each element of D serves to gather and store data reflecting the organization's operations and status.
- Assume that IT features reporting mechanisms that allow data repositories to decompose horizontally and vertically, as well as other operations ('filtering', 'rollup', 'drill-down', 'slice & dice', and 'pivot') (Chaudhuri, 1997; OLAPCouncil, 1997). This decomposition generates new data repositories that integrate into D. DM uses these mechanisms to search for helpful data to solve decision problem.

The underlying complexity of each decision-making problem in the organization lies in the need for coordinating the intelligent assessment of external factors with the effective use of required data.

4.2.1.2 Objective

The objective is to define the set of relevant data repositories for a decision-making problem, considering the state of the environmental factors (importance and value ratings). This definition also requires the assessment of the system state, which is embedded in the usefulness ratings assigned to all of the elements of D.

The second objective is to embed in the modeling methodology the functionality to capture the way decisions are made to build corporate memory. Corporate memory should assist the decision maker to improve consistency and accuracy in decisions. Consistency is achieved by recalling the way decisions were made in the past. Accuracy is achieved by facilitating the access to data previously defined as relevant.

4.2.1.3 Notation

E, set of all environmental factors (qualitative data)

= {'Material Delivery Improvement', 'Material Expedited Costs', 'Competition Price Level'…}

$E = \{e_k \in E\};\ k = 1, 2, ..., |E|$ [5]

D, set of all data repositories (quantitative data, e.g., tables, reports, charts, etc.)

= {'Demand forecast', 'Production history', 'Production Backlog'…}

---

[4] See previous subsection

[5] During the ethnographic study, the expert decision maker considered eleven environmental factors

$D = \{d_j \in D\};\ j = 1, 2, ..., |D|$[6]

I, set of all *discrete ordinal* categories used to judge *importance* of each environmental factor

$I = \{i_p \in I\};\ p = 1, 2, ..., |I|$; For the domain studied $|I| = 4$; therefore,

$I = \{i_p \in I\};\ p = (1, 2, 3, 4);\ I = \{Extreme,\ High,\ Medium,\ Low\}$

V, set of all *discrete ordinal* categories used to judge *state value* of each environmental factor

$V = \{v_q \in V\};\ q = 1, 2, ..., |V|$; for the domain studied $|V| = 3$; therefore,

$V = \{v_q \in V\};\ q = (1, 2, 3);\ V = \{Adverse,\ Neutral,\ Favorable\}$

D x E, set of all ordered pairs (D, E)

$D \times E = \{(d_j, e_k) : d_j \in D;\ e_k \in E\};\ j = 1, 2, ..., |D|,\ k = 1, 2, ..., |E|$

I x V, set of all ordered pairs (I, V)

$I \times V = \{(i_p, v_q) : i_p \in I;\ v_q \in V\};\ p = 1, 2, ..., |I|,\ q = 1, 2, ..., |V|$

N = (D, A), directed network formed by all data repositories and arcs

A, set of arcs connecting data repositories

$N_R = (D_R, A_R)$, directed network formed by all the *relevant* data repositories and the arcs; $N_R \subseteq N$

D_R, set of all data repositories that become *relevant* as a result of applying heuristic; $D_R \subseteq D$

A_R, set of arcs connecting *relevant* data repositories; $A_R \subseteq A$

[6] During the ethnographic study, we captured expert knowledge of relevance conditions for a subset of 194 data repositories (documents and reports)

4.2.1.4 The order properties of elements of the sets I and V

Based on the ethnographic study it was determined that decision makers impose a relationship of order on the elements of the sets *I* and *V*. We describe these relationships in the following definitions:

<u>Definition 1 "Properties of order of set I":</u>

The decision maker imposes a relationship of order on the elements of *I* by associating each category to its needs for specialized information, i.e., a situation ranked with 'high' importance implies needs for more information than a situation with 'low' importance ranking. Thus considering the four categories of importance decision maker used during the ethnographic study, we established properties of order on the elements of *I* as follows:

'Extreme' » 'High' » 'Medium' » 'Low'

If we associate each category to a numeric scale, i.e., i_1 = 'Extreme', i_2 = 'High', i_3 = 'Medium' and i_4 = 'Low':

$$i_1 \gg i_2 \gg i_3 \gg i_4$$

<u>Remark:</u> The symbol '»' reads *"has needs for more information than", or "has dominance over"*.

Following a similar argument, the properties of order for the elements of "**V**" are established in the following definition.

<u>Definition 2 "Properties of order of set V":</u>

The decision maker imposes a relationship of order on the elements of *V* by associating each category to its needs for specialized information, i.e., a situation judged with ‘adverse’ conditions implies needs for more information than a situation judged as ‘favorable’. Thus considering the three categories of state values the expert decision maker used during the ethnographic study, we established properties of order on the elements of *V* as follows:

‘Adverse’ » ‘Neutral’ » ‘Favorable’

Associating each category to a numeric scale, i.e., v_1 = ‘Adverse’; v_2 = ‘Neutral’; v_3 = ‘Favorable’, we have that:

$$v_1 \gg v_2 \gg v_3$$

Definition 3 “Properties of order of set $I \times V$

Consider the elements i_p, $i_r \in$ I and v_q, $v_s \in$ V. Assume that $i_p \neq i_r$, and $v_q \neq v_s$.

The properties of order on any two elements of the set $I \times V$, are established as follows:

Let (i_p, v_q) and (i_r, v_s) be any two elements of $I \times V$. We say that $(i_p, v_q) \gg (i_r, v_s)$, i.e., *“$(i_p, v_q)$ has needs for more information than (i_r, v_s)”* when its ordinal values (for both importance and state categories) are *smaller than or equal to* the corresponding values of the other element.

$$(i_p, v_q) \gg (i_r, v_s)\ if\ \ p \leq r \vee q \leq s;\ p, r \in (1, 2, \ldots, |I|);\ \ q, s \in (1, 2, \ldots, |V|)$$

The terms of this definition will be clarified with the following examples.

Example 1:

Consider the element (*3, 2*) $\in I \times V$, formed by the two components ‘medium’ and ‘neutral’[7]. »

[7] Using the numeric associations to each importance category we have that ‘Extreme’ = 1, ‘High’ = 2, ‘Medium’ = 3, and ‘Low’ = 4; similarly, the numeric associations to each state category are: ‘Adverse’ = 1, ‘Neutral’ = 2, and ‘Favorable’ = 3.

The *dominance relationships* of this element are the following (see Figure 15):

$(3, 2) \gg (3, 3)$; *because* $3 \leq 3$ *and* $2 \leq 3$

$(3, 2) \gg (4, 2)$; *because* $3 \leq 4$ *and* $2 \leq 2$

$(3, 2) \gg (4, 3)$; *because* $3 \leq 4$ *and* $2 \leq 3$

	Adverse (1)	Neutral (2)	Favorable (3)
Low (4)	(4, 1)	(4, 2)	(4, 3)
Medium (3)	(3, 1)	(3, 2)	(3, 3)
High (2)	(2, 1)	(2, 2)	(2, 3)
Extreme (1)	(1, 1)	(1, 2)	(1, 3)

Figure 15: The dominance region of the pair (3, 2) » {(3, 3), (4, 2), (4, 3)}

Example 2:

Consider now the elements (2, 2) and (3, 1). Evaluate *dominance relationships* between these two elements: $(2, 2) \overset{?}{\gg} (3,1)$.

Solution: Since $2 \leq 3$, *but* $2 \geq 1$, we conclude that none of the elements *have dominance over* the other.

4.2.1.5 Relations

$\mathbf{f_{EX}}$**:** Expert knowledge for all contexts and problems (inferred from the ethnographic study). This function defines the conditions (i_m, v_n) of environmental factor (e_k) under which a data repository (d_j) becomes relevant.

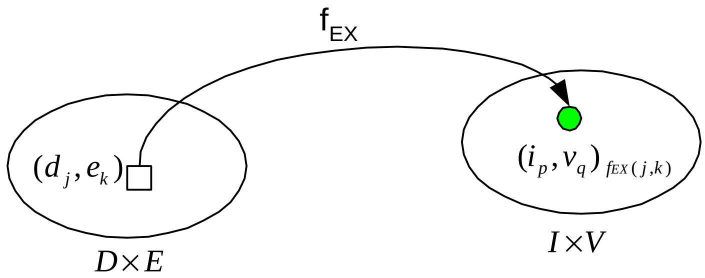


Figure 16: The f_{EX} function (Expert Knowledge)

$$f_{EX} : D \times E \rightarrow I \times V$$

$$f_{EX}(d_j, e_k) = \{ \ (i_p, v_q)_{f_{EX}(j,k)} : i_p \in I; v_q \in V \ \}$$

f_{DM}: Decision maker's assessment of the state of the environmental factor e_k for a given problem.

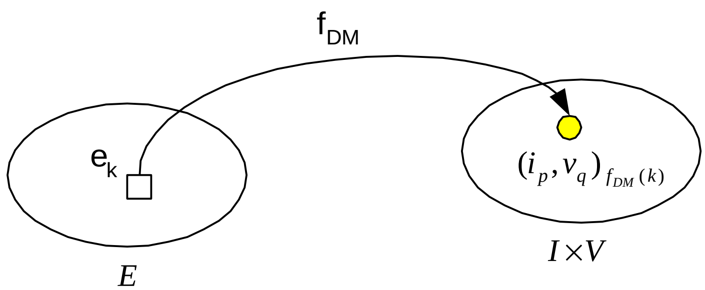


Figure 17: The f_{DM} function (Decision Maker assessment)

$$f_{DM} : E \rightarrow I \times V$$

$$f_{DM}(e_k) = \{ \ (i_p, v_q)_{f_{DM}(k)} : i_p \in I; v_q \in V \ \}$$

The definition of the properties of any environmental factor (e_k) through the function $f_{DM}(e_k)$ permits the decision maker to define a subset of data repositories $\{d_j\}$ that are relevant for the decision at hand:

$$e_k(i_p, v_q)_{f_{DM}(k)} \rightarrow \{ d_j \}$$

More details about this process are provided in the next subsection.

4.2.2 Heuristic algorithm

The objective of the heuristic algorithm is to determine whether a data repository becomes relevant for the decision at hand. For that, the algorithm uses two inputs, the first is the information (expert knowledge) of the conditions under which each data repository becomes relevant; the second is the evaluation of current conditions (decision maker's assessment). With these two inputs, and the dominance relationships embedded the heuristic algorithm determines whether or not a certain data repository becomes relevant for the decision at hand.

The heuristic algorithm uses the relations f_{EX} and f_{DM}, as well as the properties of order established in Definitions 1, 2, and 3. An overview of the heuristic algorithm is provided in the pseudo code depicted in Figure 18.

Heuristic Algorithm: Find relevancy of all d_j in D

$\{for\ \ j = 1\ to\ \ j = |D|\};\ |D| = 194$

Do

$\{for\ k = 1\ to\ k = |E|\};\ |E| = 11$

Do

$f_{EX}(d_j, e_k) = \{\ (i_p, v_q)_{f_{EX}(j,k)} \in X : i_p \in I;\ v_q \in V\ \};$ *Expert knowledge*

$f_{DM}(e_k) = \{\ (i_p, v_q)_{f_{DM}(k)} \in X : i_p \in I; v_q \in V\ \};$ *DM's assessment*

$if\ \left(f_{DM}(e_k) \gg f_{EX}(d_j, e_k)\right)*$

$then \begin{cases} (d_j \in N_R | e_k) \\ Next\ \ j \end{cases}$

$else \begin{cases} (d_j \notin N_R | e_k) \\ Next\ k \end{cases}$

endIf

endOfFor k

endOfFor j

Where: $I = \{i_p \in I\};\ p \in (1, 2, 3, 4);\ I = \{Extreme, High, Medium, Low\}$

$V = \{v_q \in V\};\ q \in (1, 2, 3);\ V = \{Adverse, Neutral, Favorable\}$

$I \times V = \{(i_p, v_q) : i_p \in I;\ v_q \in V\};\ p = 1, 2, 3, 4;\quad q = 1, 2, 3$

(*) Note:

$if\ \left(f_{DM}(e_k) \gg f_{EX}(d_j, e_k)\right)$

i.e., current conditions are worse (more adverse and more important) than historic (knowledge base). Therefore, document d_j is still relevant

Figure 18: Pseudo code of heuristic algorithm

4.3 Summary

In this chapter we presented a modeling methodology to capture and structure the expertise required when dealing with qualitative and quantitative data in decision problems. In the model, expertise is characterized into three separate processes: (i) assessment of the qualitative data, in the form of ordinal quantities or an ordered representation (ii) associating the values assessed by the expert for qualitative data with specific quantitative data; and finally (iii) structuring the information space necessary to search and access the required quantitative data. Figure 19 depicts the processes that model expertise. Details are explained next.

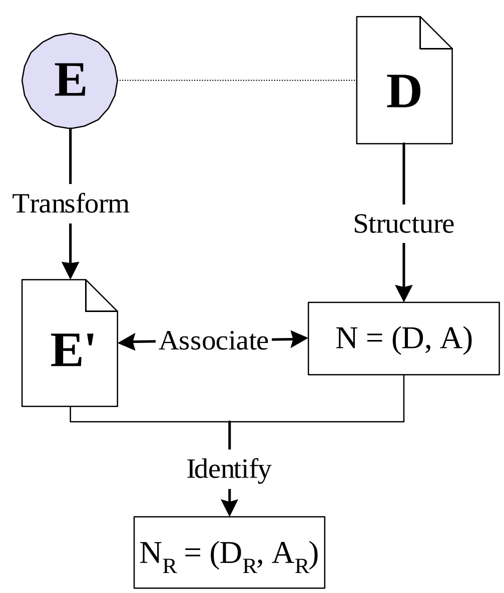


Figure 19: Process model of expertise.

The process of *transforming* qualitative data into a quantitative assessment implies the representation of the influencing external factors and their attributes within an ordinal scale or an ordered representation. The following procedure is proposed. First, all external factors (E) must be identified. Then, a set of attributes describing each factor is defined; for instance, if ‘Market growth’ is an external factor, one possible attribute of it could be ‘current growth tendency’. It is not necessary that all the factors possess the same set of attributes. Finally, an ordinal measurement scale over each attribute is proposed. This measurement system is used to represent any particular configuration of external factors and their attributes at any point in time; later, this configuration is mapped into the needed quantitative data.

For the process of *associating* assessed values of qualitative data (E') with quantitative data represented as a graph (N), a model for a decision aid is proposed. The aid incorporates not only a set of rules, but also cognitive considerations that permit mapping any particular configuration of *transformed* qualitative data into a sub-graph containing only needed quantitative data (N_R). In this chapter we referred to the decision aid as the heuristic algorithm, which uses dominance relationships to determine the set of needed quantitative data.

Finally, for the process of *structuring* the information space necessary to search and access the required quantitative data (N_R), the goal is to identify both the data elements and the sequence in which they are accessed. The identification process is exhaustive, i.e., the model requires capturing and structuring all possible data elements that an expert decision maker would use for as many circumstances as possible. The methodology uses a directed graph structure (N) to represent the relationships between data and the sequence in which an expert created or accessed them.

In this chapter we developed an instance of the modeling methodology and the heuristic algorithm. Elements are depicted in Figure 20.

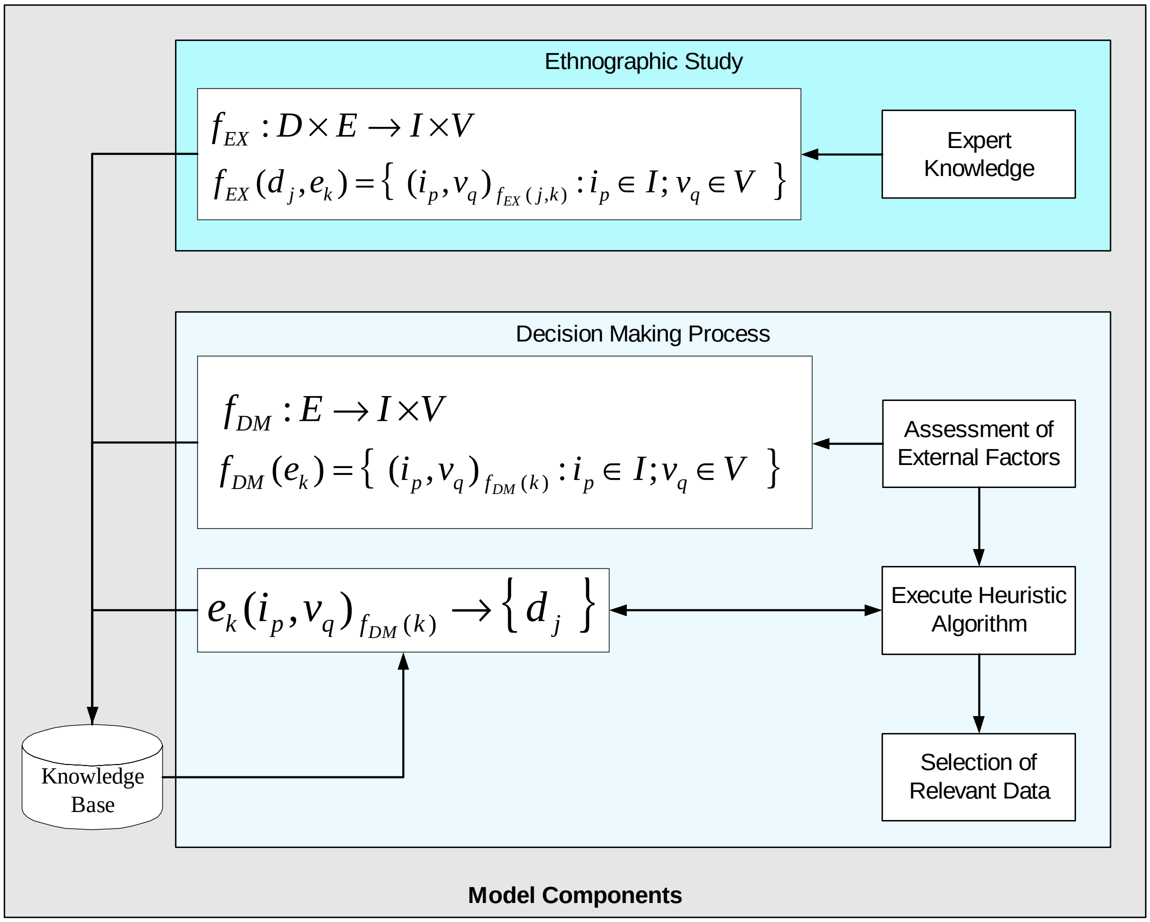


Figure 20: Implementation of the modeling methodology and heuristic algorithm

The ethnographic study (presented in Chapter 3), provided the empirical evidence for capturing the expertise required to for mapping *transformed* qualitative data (E') into quantitative data represented as a graph. In this chapter we referred to this mapping as the f_{EX} relationship. We also presented the rules for *transforming* qualitative data into a quantitative assessment. For this, we represented the influencing external factors and their attributes within an ordinal scale or an ordered representation. The process of assessing a particular configuration of attributes for the set of external factors was referred to as the f_{DM} relationship. Finally, the particular implementation of the algorithm receives the two relationships (f_{EX} and f_{DM}) as inputs and uses the dominance relationships (properties of order) to determine the set of required data repositories for the decision problem. A computer-based implementation of the elements of the modeling methodology and heuristic algorithm depicted in Figure 20 is described in the next chapter.

CHAPTER 5

SOFTWARE IMPLEMENTATION

In this chapter, we describe a computational implementation of the graph-based modeling methodology to support decision aiding. The computational implementation incorporates a mechanism to allow the decision maker to assess the conditions of external factors. This assessment provides a structure for multiple types of qualitative data, and then leads to a search of relevant pieces of quantitative data normally stored in databases provided by the IT applications (e.g., ERP implementations). The connections and relationships between the assessment of environmental factors (qualitative data) and data stored in information management systems (quantitative data) are presented to the decision maker through an interactive graph-based interface, having features that permit the decision maker to access relevant pieces of information.

As a proof-of-concept application, we are addressing the decisions associated with production planning in a manufacturing organization. These decisions address the problem of “what to produce”, “when to produce”, and “from where/whom to acquire the required resources”.

Although the computational implementation is domain dependent, we believe that the underlying model can be easily adapted to other domains where decision makers deal with these two types of data: data originating from external factors (e.g., Market Trends, political issues, economical forces, external decision makers, etc) and data warehouses.

5.1 Architecture of the computational prototype

The computational implementation requires two high-level components: (i) a prototypical ERP implementation to show user-interaction features (*miniERP* transactional interface), and (ii) a graph-based decision support system (*miniERP-GDSS*), which implements the modeling approach. Figure 21 depicts the high-level components.

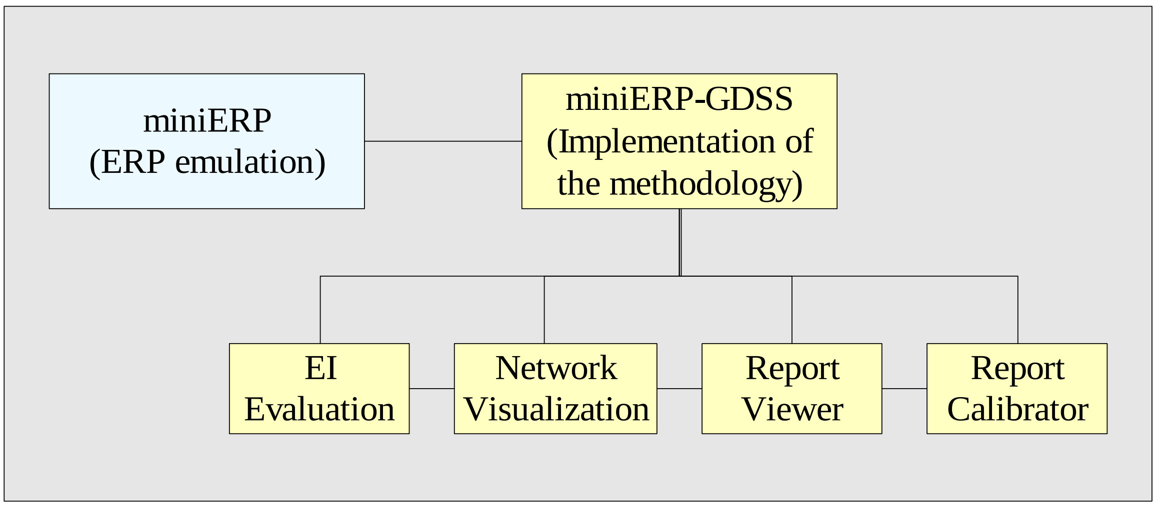
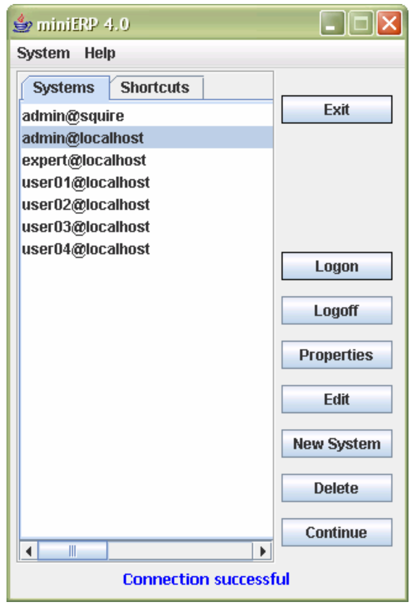


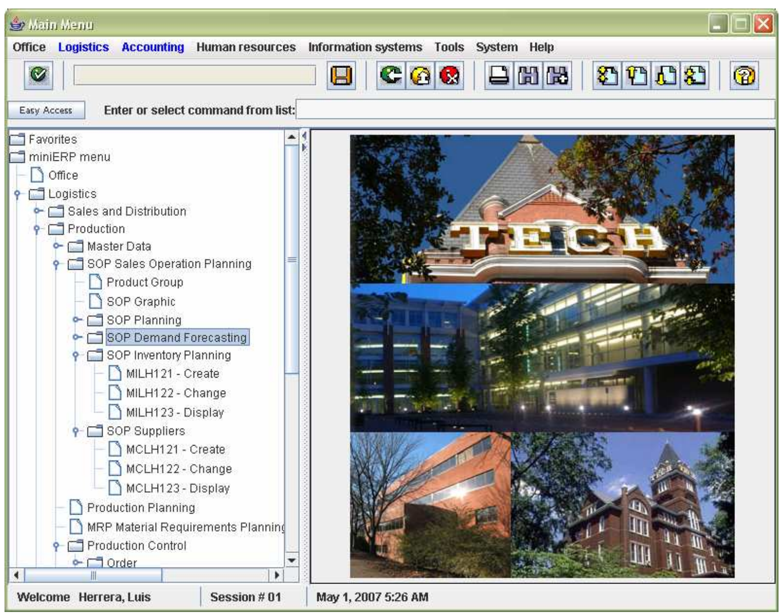
Figure 21: *miniERP* – Computational Implementation

5.1.1 Component A: *miniERP* transactional interface

We have developed *miniERP*, a simplified emulation of typical commercial ERP implementations. Even though, it is not "commercial strength", and still lacks full functionality, it incorporates most basic features of an ERP application, e.g., various navigation features – tree menu, command line, user authentication, bookmarks, typical interfaces to query information system, and communication to a spreadsheet for data management. Figure 22 depicts some displays of the *miniERP* transactional interface.



(a) *miniERP* - Logon Interface



(b) *miniERP* – Main Menu Interface

Figure 22: *miniERP* – Transactional Interface

The objective of the *miniERP* transactional interface is to show how decision makers normally interact with an ERP implementation to make decisions. The *miniERP* transactional interface will serve as baseline to compare the benefits achieved with the computational implementation of the network-based methodology (*miniERP-GDSS*), as described next.

5.1.2 Component B: *miniERP-GDSS* decision support interface

The *miniERP-GDSS* decision support interface consists of four components: (a) the *External Factors (EI) Evaluation* interface to assess external factors, (b) the *Network Visualization* interface to represent and visualize data and their relationships as a network, and also to execute first-level analysis (data clustering), (c) the *Report Viewer* interface permits the visualization of accessed data for further analysis; finally (d) the *Report Calibrator* interface to define usefulness rating to report. This interface permits the visualization of accessed data for further analysis. Figure 21 depicts high-level components of computational prototype.

5.2 Design of the *miniERP-GDSS*

In this subsection we describe in detail each component of the *miniERP* Graph-based Decision Support System (*miniERP-GDSS*). The *miniERP-GDSS* is comprised of four components: (i) *EI Evaluation*, (ii) *Network Visualization*, (iii) *Report Viewer*, and (iv) *Report Calibrator*. The access to these components is controlled from the *Main Menu* interface (see Figure 23). The function of each component is described next.

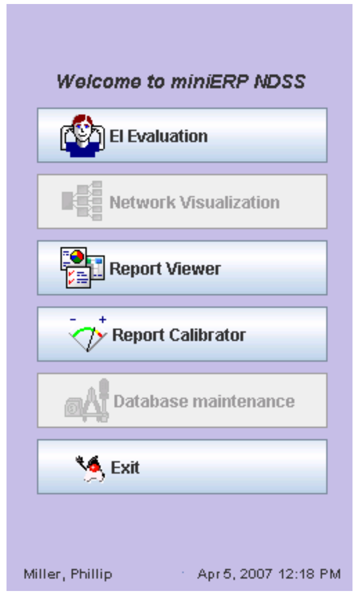


Figure 23: *miniERP-GDSS* – Main Menu Interface

5.2.1 The *EI Evaluation* Interface

The *EI Evaluation* component allows the decision maker to describe a specific configuration of environmental variables. Figure 24 depicts the *EI Evaluation* interface. The decision maker assesses external factors (E) in two forms: (1) I, the relative importance rating of each E – this is a subjective assessment based on the user's experience and business knowledge, and (2) V, a state value of the E at hand. For instance, suppose the E in focus is Competition Delivery Time. For this, the decision maker could say: (1) Importance level = "extremely important" and (2) Current State value = "long delivery times".

The computational prototype incorporates the heuristic algorithm (defined in previous section) that reads two inputs (i) particular configuration of environmental factors (I, V); and (ii) the usefulness values of each data repository (D). These values are stored in the form of tables in a database. The prototype reads all the inputs, and using the heuristic algorithm computes a network of relevant data. The next component serves to display a network of relevant nodes.

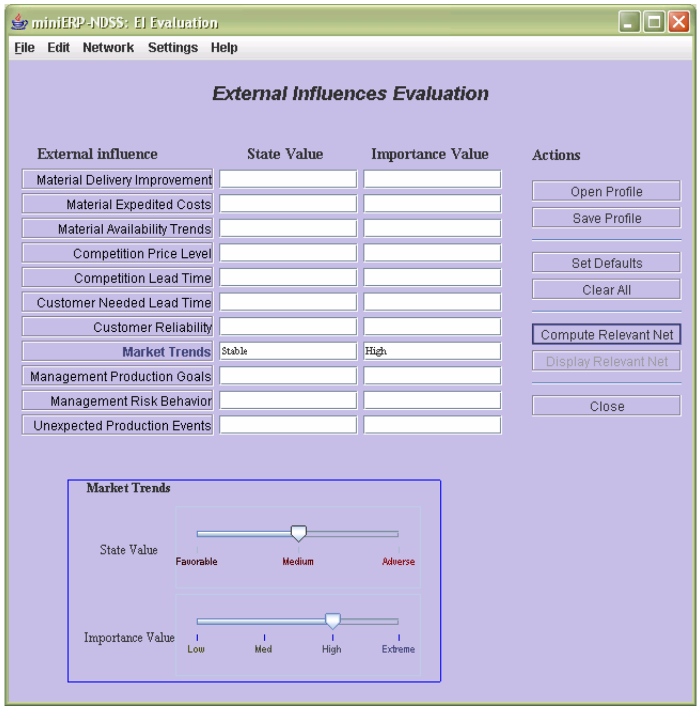


Figure 24: *miniERP-GDSS – EI Evaluation* Interface

5.2.2 The *Network Visualization* Interface

Once a particular network is created, a *Network Visualization* interface enriched with clustering and exploration features allows a user to perform first-level analysis. This analysis is possible due to the features included in the interface, e.g., search, community formation, comparisons between nodes, etc. All these features allow the user to examine multiple relations between nodes for any given configuration of external factors.

The characteristics of all the nodes are stored in a database and read from there to create a particular network. Decision makers can visually discover those nodes of interests by utilizing the various filtering and search tools included in the visualization display. Once the decision maker inputs an initial configuration of variables, the program generates clusters of vertices (of documents) that are categorized by groups of different interest. The user can also manipulate the emphasis and increase or decrease the degree of accuracy. Figure 25 depicts the *Network Visualization* interface.

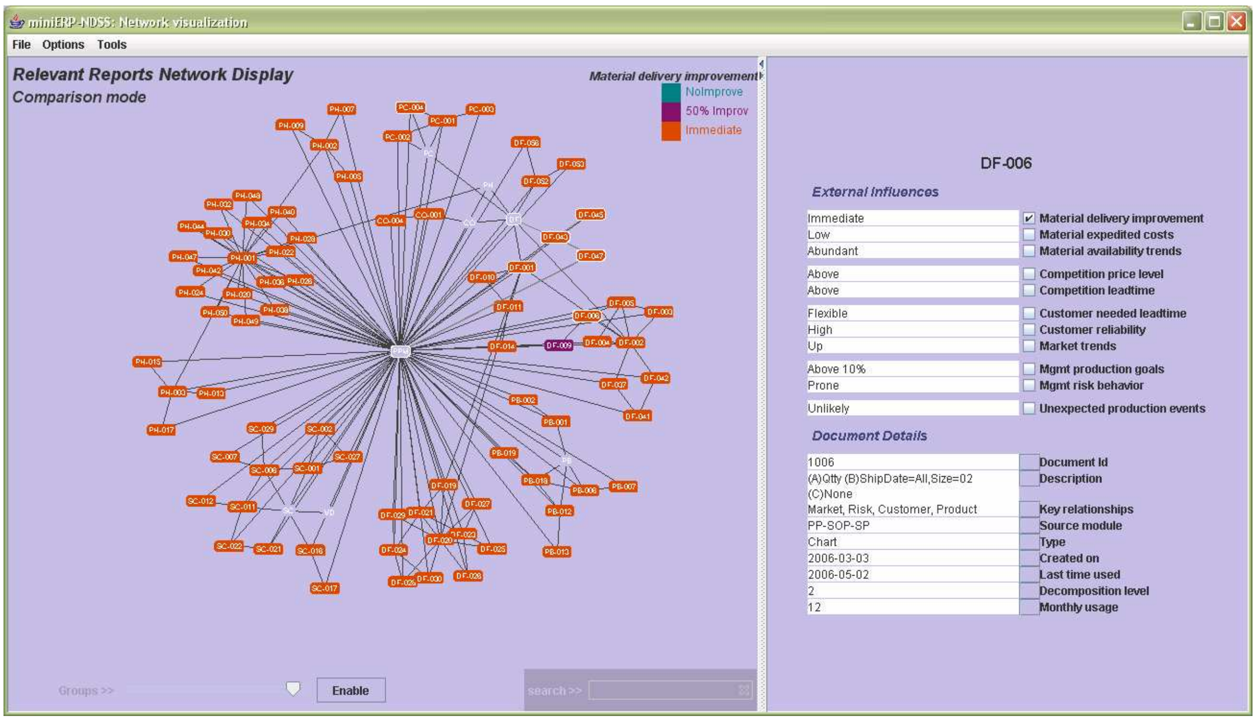


Figure 25: *miniERP-GDSS – Network Visualization* Interface

5.2.3 <u>The *Report Viewer* Interface</u>

The Report Viewer component provides a workspace area for manipulating and analyzing documents marked as relevant. In this space, the decision maker can study special details for the problem at hand. Figure 26 depicts the *Report Viewer* interface.

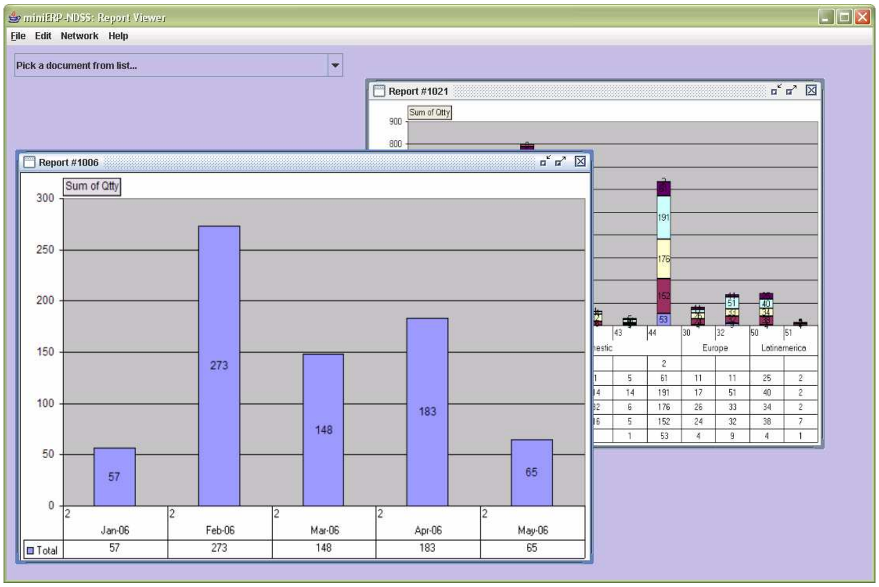


Figure 26: *miniERP-GDSS – Report Viewer* Interface

5.2.4 <u>The *Report Calibrator* Interface</u>

Every time a document is incorporated in the application database (either extracted from the data warehouse or created as a result of a business application analysis) it is defined with certain attributes that will permit the application to include it under certain circumstances (a problem assessment). This process was defined as "usefulness assessment" in the previous section. The objective of the *Report Calibrator* interface is to allow a user to assign a usefulness rating to each report / data repository with respect to each environmental factor.

5.3 Implementation details

5.3.1 <u>Clustering algorithms</u>

A common technique for performing data analysis in networks is based on data clustering, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is the classification of similar objects into groups, or more precisely, the partitioning of a data set into subsets (clusters), so that the data in each subset (ideally) share some common trait - often proximity according to some defined distance measure. Our implementation uses of two algorithms to study formation of clusters in the network: (i) the Spring-embedded algorithm (Anonymous, 2007a); and (ii) an extension of the Girvan and Newman algorithm to find community structure in networks (Newman & Girvan, 2004).

5.3.2 Visualization techniques

The visualization of the network is built using Java 2D. Two Open Source visualization toolkits are used: JUNG (Anonymous, 2007a) and Prefuse (Anonymous, 2007b).

5.4 Summary

In this chapter we presented details of the computational implementation of the graph-based modeling methodology to support decision aiding. The computational implementation comprises two high-level components: (i) a prototypical ERP implementation to show user-interaction features (*miniERP* transactional interface), and (ii) a graph-based decision support system (*miniERP-GDSS* decision support interface), which implements the modeling methodology and the heuristic algorithm presented in Chapter 4.

The first component, *miniERP* transactional interface, has been developed to demonstrate interactions with a prototypical ERP system. This system will serve as a control system during the empirical evaluation of the model (Chapter 6).

The second component, *miniERP-GDSS* decision support interface incorporates various functions. The first function allows the decision maker to assess the conditions of environmental factors. This assessment provides a structure for multiple types of qualitative data, and then leads to a search of relevant pieces of quantitative data normally stored in databases provided by the IT applications. The second function, an interactive graph-based interface provides a visualization of the data stored in information systems (quantitative data) and their relationships with different attributes of environmental factors (qualitative data). The graph-based interface allows users to narrow search for data and to access specific reports (third function). Finally, the *miniERP-GDSS* incorporates a function that allows decision makers to assess or change the relationships between qualitative and quantitative data.

An empirical evaluation of the software implementation is presented in the chapter that follows. The two types of interfaces *miniERP* and *miniERP-GDSS* will serve as control and experimental conditions respectively. Our intention is to test the benefits of the modeling methodology and the heuristic algorithm to aid in the decision making process.

.

CHAPTER 6

EMPIRICAL EVALUATION

An empirical evaluation of the Graph-based Decision Support System proof-of-concept *miniERP-GDSS* Assistant is presented in this chapter. The evaluation assesses the effectiveness of the Software Assistant in supporting production planners during the creation of a rough-cut material requirements plan (RC-MRP).

This chapter is organized as follows: in the first section we describe the method used to evaluate the software, participants, procedure and details of the task to be tested. Next, we present a detailed description of the experimental design, factors, levels, and the set of performance measures.

6.1 Method

6.1.1 Participants

An expert decision maker and three production planners (non-expert users) certified in the use of the production planning module of the implemented ERP system at the Cooper Cameron Valves plant participated in the evaluation. Profiles of participants are provided in Appendix C.

6.1.2 Procedure

The production planners participated in two free-time sessions. The duration of each session ranged from two to four hours. In one session participants were asked to solve three regular problems of rough-cut material requirements plan utilizing the transactional interface of *miniERP* (control condition). The environmental circumstances of each problem varied, which resulted in different levels of difficulty for the problems. Details of problems description and their characteristics are provided in Appendix D. In another session, the participants were asked to solve the set of three problems, but utilizing the *miniERP-GDSS* Decision Support Interface (experimental condition).

The participants completed one session per day. Within a session, they were granted short breaks between each problem. Due to their normal work activities, participants were allowed to take longer breaks (more than 60 minutes) or even miss a day between problems corresponding to a single session. Table 9 summarizes the contents and duration of the two problem-solving sessions.

Table 9: Overview of RC-MRP problem solving sessions

Session	Purpose	Activities	Duration
1	Test problem solving performance and accuracy with standard (transactional) interface	• Introduction	3 minutes
		• System description and explanation	10 minutes
		• Lesson A (*): *miniERP* Transactional interface training	20 minutes
		• Instructions for problem #1	5 minutes
		• Problem #1 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #2	5 minutes
		• Problem #2 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #3	5 minutes
		• Problem #3 – timed RC-MRP execution	free
2	Test problem solving performance and accuracy with decision support system interface (*miniERP-GDSS*)	• Introduction	3 minutes
		• System description and explanation	10 minutes
		• Lesson B (*): *miniERP-GDSS* Decision Support Interface training	20 minutes
		• Instructions for problem #1	5 minutes
		• Problem #1 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #2	5 minutes
		• Problem #2 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #3	5 minutes
		• Problem #3 – timed RC-MRP execution	Free

(*) Prior to each problem solving session, participants received two training lessons (A and B). These lessons permitted participants become familiar with the interfaces (transactional and decision support) required to solve the problems.

6.1.3 Required output

Participants were asked to solve six regular rough-cut material requirements plan (RC-MRP) problems. These problems were presented in two sets. Table 10 displays a sample output of the RC-MRP problem. This table is composed of two entries for each period and each product. An entry comprises of a pair of values per period for each row. The first value (qualitative) refers to the material supplier. The second one is a quantitative value and it refers to the amount of products to be manufactured for each period.

Table 10: Sample output of the RC-MRP problem

Product size	Product family	Supplier (code) | Quantities (units)													
		Oct-01		Nov-01		Dec-01		Jan-02		Feb-02		Mar-02	
		Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty
2	02" - 06"	MSup-43573	103	MSup-43573	107	MSup-43573	111	MSup-43573	111	MSup-43573	112	MSup-43573	116
3	02" - 06"	MSup-43573	110	MSup-43573	112	MSup-43573	112	MSup-43573	117	MSup-43573	119	MSup-43573	123
4	02" - 06"	MSup-43573	110	MSup-43573	112	MSup-43573	112	MSup-43573	117	MSup-43573	119	MSup-43573	123
6	02" - 06"	MSup-43573	276	MSup-43573	282	MSup-43573	289	MSup-43573	295	MSup-43573	302	MSup-43573	308
8	08" - 12"	MSup-25789	52	MSup-25789	52	MSup-25789	52	MSup-25789	56	MSup-25789	56	MSup-25789	56
10	08" - 12"	MSup-25789	48	MSup-25789	48	MSup-25789	48	MSup-25789	52	MSup-25789	52	MSup-25789	52
12	08" - 12"	MSup-25789	48	MSup-25789	48	MSup-25789	48	MSup-25789	44	MSup-25789	44	MSup-25789	44
16	14" - larger	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53	MSup-43544	53
20	14" - larger	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50	MSup-43544	50
24	14" - larger	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31
30	14" - larger	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38	MSup-13170	38
36	14" - larger	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31	MSup-13170	31

6.2 Experimental design

An experimental study was completed to evaluate the benefits of the *miniERP-GDSS* interface. A multi-factor analysis of variance model was used to analyze the marginal effects and interaction of the experimental factors. Three factors (explanatory or predictor variables) were studied: (i) inclusion or no-inclusion of graph-based model features (*miniERP-GDSS*), (ii) expertise levels of participants, and (iii) difficulty level of exercises. Within each factor, several levels were considered. The high-level of expertise required for the subjects participating in the evaluation imposes constraints on the number of participants available. This situation forced the experiment to have one case per treatment; this situation has been considered in the statistical analysis. Additionally, the subjects participating in the event performed in all the factors. This situation is of special interest for the first factor mentioned above "inclusion or no-inclusion of design features" due to the appearance of carry-over effects. Finally, the responses of interest are grouped into two categories: (i) benefits to access the required data for the problem at hand, and (ii) benefits to achieve better solutions. Details of factors, treatments, and performance measures are provided next.

6.2.1 Experimental factors and levels

A multi-factor analysis of variance model is proposed to analyze the marginal effects and interaction of the experimental factors. Three factors (explanatory or predictor variables) were studied: (i) interface type, i.e., inclusion or no-inclusion of graph-based model features (graph-based model), (ii) difficulty level of exercises, and (iii) expertise levels of participants.

6.2.1.1 Interface type

Interface type refers to the inclusion or no-inclusion of the design features. The design features correspond to the software implementation of the graph-based model discussed in the previous two chapters of this document. Therefore, two levels (conditions) for this factor were considered: (1) Control and (2) Experimental. Control condition (C) corresponds to solving the set of three RC-MRP problems using the *miniERP* interface, which lacks the design features. The *miniERP* interface is a somewhat simplified computer-based implementation of the regular ERP system that participants use in their daily work environment. For a detailed description of the *miniERP* interface refer to Chapter 5.

In the experimental condition (E), subjects were required to solve the set of RC-MRP problems using the *miniERP-GDSS* interface that includes the design features. These features permit the access to relevant data through the assessment of the importance and state value of external factors. These features have some effect on two types of responses of interest: (1) they facilitate the access to relevant data for the decision-making process, and (2) they aid the decision makers to achieve better solutions for the decision-making problems. More details about these effects are provided in a subsequent subsection.

6.2.1.2 Difficulty level of exercises

Subjects were asked to solve three problems of rough-cut material requirements plan. The problems varied in difficulty level: (1) favorable, (2), adverse, and (3) mixed. The difficulty level of each problem was determined by the characteristics of the state value of environmental factors[8]. For instance, in the first problem, it is assumed that conditions of the environmental factors are all favorable, e.g., market trends are growing, delivery times from raw material vendors are short, etc. Opposite circumstances of environmental factors characterize the second problem. For instance, customers' required production times are very short and not negotiable. Finally, in the third exercise, a mix of circumstances is presented, i.e., some environmental factors present adverse conditions,

[8] Environmental factors present three possible levels of state values: (1) Adverse, (2) Neutral, and (3) Favorable. See details in Chapter 4.

whereas others present favorable, and the rest have neutral conditions. For a detailed description of the problems, refer to Appendix D.

6.2.1.3 Expertise level of subjects

Four levels of expertise are being considered one for each of the four subjects participating in this experimental study: (1) Extreme, (2) High, (3) Medium, and (4) Low. The subject possessing the extreme level of expertise is known as the “Expert” subject, the others are called as “Non-expert” subjects. The discrimination of subjects into these two categories is used to describe the performance of non-expert decision makers. Profiles of subjects describing their expertise levels and roles in the production planning process at the CCV Plant are described in Appendix B.

6.2.2 Experimental treatments and repetitions

In multi-factor analysis of variance models, with factors having several levels, a treatment corresponds to a combination of factor levels. For the proposed experimental study we have the following factors and levels:

- Interface type (experimental factor): two levels
- Difficulty level (experimental factor): three levels
- Expertise level (classification factor): four levels

Therefore, the experimental study has a total of 24 treatments (=2x3x4).

The limited availability of subjects with a minimum expertise for solving the problem at hand (RC-MRP) severely limited the number of observations that could be obtained. Therefore, only one replication of the experiment was possible for each treatment. That is, each of the four subjects performed on each of the levels of the other two factors, i.e., each of the three problems (‘favorable’, ‘adverse’ and ‘mixed’) using both levels ('C' and 'E'). Consequently, the experimental study yields 24 output forms (solutions) for the RC-MRP problems. For each output (solution) researcher recorded all the transactions executed, e.g., access to data, data transformations, execution times, etc.; analyses of both types of output forms for each treatment are considered in the analysis of performance measures.

A modification of the ANOVA model was required for the analysis of the three-factor study due to the single replication per treatment. More details of the ANOVA model formulation and results are provided in Chapter 7.

6.2.3 Carry-over effects

Subjects performed each level of the 'Interface type' factor in one session. In order to consider the carry-over effects, half of the subjects performed in level C (control condition) first and then in level E (experimental condition); the other subjects performed in level E first and then in level C. Within each condition, subjects solved three problems of rough-cut material requirements plan. Table 11 shows the order in which subjects solved the problems in both conditions.

Table 11: Order in which subjects solved the problems in both conditions

Subject	First session			Second session		
Subject #1	E3	E2	E1	C2	C3	C1
Subject #2	E2	E1	E3	C1	C2	C3
Subject #3	C3	C1	C2	E1	E2	E3
Subject #4	C3	C2	C1	E1	E3	E2

6.2.4 Dependent variables (measures of performance)

A total of 24 treatments comprise this experimental study. A comparison of the marginal and interaction effects are studied here. Measures of performance recorded for each one of the 24 treatments are grouped into two categories: (i) *transactional* performance measures, i.e., benefits to access the required data for the problem at hand, and (ii) *numeric* performance measures. An explanation and further details of each dependent variable are provided next.

6.2.4.1 Transactional performance measures

These measures are aimed to analyze the benefits of the experimental condition (E) for accessing the required data for the problem solving process. Three performance measures are grouped into this category. The first one “Execution time” reflects directly the improvements for accessing required data to solve the problems at hand. The second and third performance measures are used to evaluate the model. Table 12 depicts the performance measures considered in this category and the variables recorded. Details of these measures follow.

Table 12: Transactional performance measures

M1: Execution time
M1.1. Total time required to compute a solution, A=B+C.
M1.2. Time spent **creating/browsing** documents.
M1.3. Time spent **using** relevant documents to compute a solution.
M1.4. Ratio: (M1.2) / (M1.1): the percentage of time spent creating/browsing documents.
M2: Data overload level
M2.1. Number of documents **created/browsed.**
M2.2. Number of documents **used** to compute a solution.
M2.3. Ratio: (M2.1) / (M2.2): a measure of the unnecessary amount of documents **created/browsed** during the search of data. A ratio closer to 1 indicates that decision maker is dealing with smaller amount of unnecessary documents
M3: Completeness of 'Expert Knowledge'
M3.1. Number of documents **created** & **used** in control condition, but not available for **browsing** in experimental condition, i.e., missing in (KB).
M3.2. Ratio: (M3.1) / (M2.2): a measure of the completeness of the knowledge base. This ratio represents a normalization of M3.1 with respect to the number of documents used (M2.2).

M1: Execution time

Execution time corresponds to the time the decision maker took to create an output for the RC-MRP problem. Three variables were recorded: (M1.1) Total time to complete a solution; (M1.2) Time spent *creating/browsing* documents; and (M1.3) Time spent *using* relevant documents to compute a solution. Using recoded variables (M1.1) and (M1.2) a fourth variable is defined, (M1.4) the percentage of time spent creating/browsing documents, i.e., the ratio of (M1.2) / (M1.1).

The process of gathering the required data for solving the problems is depicted in Figure 27. This process comprises two phases, *create* (subject is using the control condition C - *miniERP*) or *browse* (subject is using the experimental condition E –

miniERP-GDSS), and *use*. Many documents created and browsed are not actually used for the problem solving phase. During the 'using' phase, decision maker performs offline analysis of documents. This is a cognitive process and is not described in this research. Our sole interest is in recording the time spent performing the offline analysis.

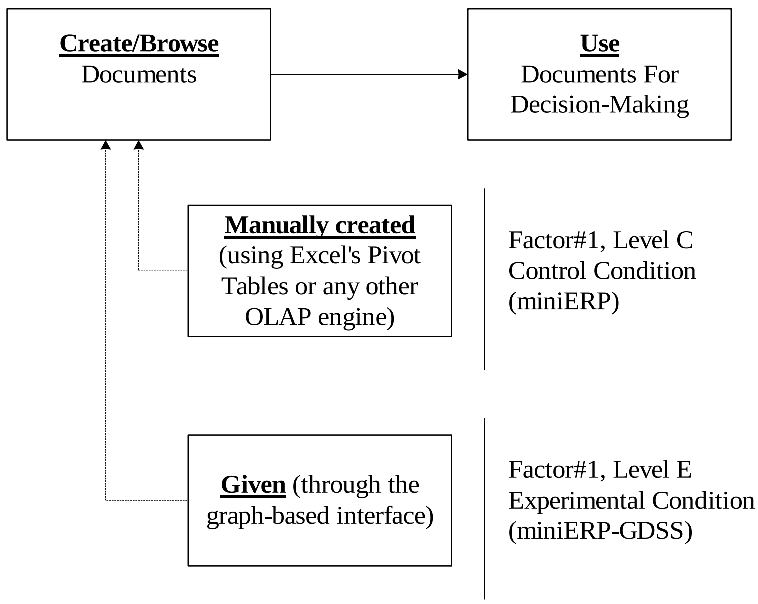


Figure 27: Process of gathering data for solving the problems

<u>M2: Data overload level</u>

When decision makers are solving the RC-MRP problem, they search for data stored in the IT system. The search process involves accessing entire 'generic' documents e.g., demand forecast for a certain period, production capacity report, etc. Decision makers manipulate these 'generic' documents to *create* 'summary' reports. In a regular RC-MRP problem-solving process this is a trial-error process. Many of these 'summary' reports are relevant and are actually used for making decisions, many others are discarded; however, the trial-error process normally leads to data overload, because of the excessive amount of unnecessary documents that are created. The regular process for accessing required data has been included in the Control condition C (level 1 of the experimental factor #1 – refer to subsection 6.2.1).

An alternate methodology to search required data is provided. The experimental interface that helps decision makers search data for specific conditions of problems is used for comparison. In the experimental interface, decision makers are not to access generic documents and create summary reports. Instead they are granted access to a set of previously created documents that match the characteristics of the problem at hand. A graph-based visualization device provides the interface so that decision makers can *browse* and select relevant documents. The experimental interface serves as the Experimental condition E, (level 2 of the experimental factor #1 – refer to subsection refer to subsection 6.2.1).

The set of recorded measures under “Data overload level” category are intended to compare the level at which the experimental factor #1, interface type (“inclusion or no-inclusion of design features – *miniERP-GDSS*) alleviates the data overload problem. During the experimental study, the researcher recorded all the processes followed by subjects to solve the sets of RC-MRP problems. Appendix F provides details of the transactions and variables recorded during the problem-solving process of one subject.

The performance measure M2 “Data overload level” provides a measure of the unnecessary documents that decision maker creates/browses to solve a RC-MRP problem. It is important to note that the number of documents created/browsed does not correspond necessarily with the number of documents that he uses, i.e., that are relevant, because many of the documents that he creates are not entirely relevant. Two variables are recorded (see Table 12): (M2.1) Number of documents created/browsed to compute a solution; and (M2.2) Number of documents used to compute a solution. Using recoded variables (M2.1) and (M2.2) a third variable is defined, (M2.3) the ratio of number of documents *created/browsed* over the number of documents actually *used*. If the value of this ratio is close to one, it indicates that most documents created/browsed resulted useful for computing a solution.

The transactional results of the experimental study will be used to compare the level at which this performance measure is achieved by the two levels of experimental factor A "Interface Type" (control level, standard *miniERP* interface and experimental level, *miniERP-GDSS* interface with the inclusion of the design features).

M3: Completeness of Expert Knowledge

Performance measure M2 intends to provide a measure of the number of documents that decision maker needs to consult to create a solution, which is an indirect measure of the data overload problem. The M3 performance measure offers some insight of the completeness of the "Expert Knowledge" (collected during the ethnographic study). For this, one variable (M3.1) is recorded during the transactional process of solving the RC-MRP problem (see Table 12). The M3.1 variable records the number of documents *created and used* in the Control Condition C (*miniERP* interface) – treatments with level 1 for factor #1 – but not available for *browsing* in Experimental Condition E (*miniERP-GDSS* interface) – treatments with level 2 for factor #1. If a document is created and used during the control condition, it simply indicated that such a document was relevant and should have been recognized during the ethnographic study and included in the "Expert knowledge" database. Clearly, a small number of documents left out during the ethnographic study indicate the completeness level of “Expert knowledge” database.

6.2.4.2 Numeric performance measures

As mentioned above (section 6.2.2), the output (RC-MRP solution) generated in each treatment required two types of analysis. The methodology for analyzing the transactional results was discussed in the previous subsection. Here we present the methodology for analyzing the numeric results of each RC-MRP solution generated for the experimental study.

A cost-based performance measure is computed for the RC-MRP generated under each treatment, i.e., for all subjects (4), all exercises (3), and interface conditions (2). This measure describes numerically the accuracy of the solutions obtained for all the treatments during the experimental study. The cost-based performance measure decomposes into three categories of cost: (1) Holding costs, (2) Acquisition costs, and (3) Opportunity cost. A description of the methodology to compute these costs follows.

<u>M4: Holding cost</u>

As described in Appendix E, the solution of a RC-MRP problem consists of the estimation of the quantities and quality of required material to satisfy forecasted demand. Once the decision maker finishes the RC-MRP, the system initiates several processes, e.g., the purchase of raw materials from the specified qualities, and the manufacture of subassemblies for the specified products.

A few periods (months) later when the actual demand occurs, the company finishes the manufacture of products and ships them to customers. This process however, is not free of deviations. For certain periods and products there will be differences in quantities between what was planned and produced (RC-MRP) and the actual demand. For any given period, when the difference between planned productions minus actual demand is positive, the company will keep an inventory of materials. The cost of carrying items in inventory includes the opportunity cost of the money invested, the expenses incurred in running a warehouse, handling costs, deterioration, damage, insurance, taxes, etc.

The computation of holding costs for the RC-MRP solutions required a reference of the actual demand; since the three cases included in the experimental study are fictitious, no actual demand was available. In order to overcome this situation, the following strategy was adopted. The output generated by the subject #1 (expert user) was taken as the reference with regard to the qualities and quantities specified for each period and product. Therefore, the output created by subject #1 (expert decision maker) was considered as the actual demand. Quantities and qualities of outputs created by other subjects were compared against the reference demand. Table 13 describes this process for treatment #C1-2 (Condition: "*Control*", Case: "*Favorable*", Subject: "*High expertise*").

For all the treatments (except for those where subject has the extreme expertise – expert decision maker), holding costs are computed as follows:

$$I_k^j = n_k^j \times r \times \left[v^j \times \left(1 - d_k^j\right)\right]$$

Where:

I_k^j = Holding cost for period k, for product j

n_k^j = Carrying inventory for period k, for product j

$n_k^j = p_k^j - q_k^j$; If $p_k^j - q_k^j \geq 0$;

$n_k^j = 0$; Otherwise

p_k^j = Forecasted quantities of product j, for period k

q_k^j = Actual demand of product j, for period k

r = Carrying cost (refer to Appendix G)

v^j = Acquisition cost for product j (refer to Appendix G)

d_k^j = Acquisition discount for raw material j during period k,

Table 13: Computation of holding costs for treatment: *C1-2*, product size: *2in*

Product size	Quantities (units) / Discount (%)											
	Nov-05		Dec-05		Jan-06		Feb-06		Mar-06		Apr-06	
	Subject#1	Subject#2	Subject#1	Subject#2	Subject#1	Subject#2	Subject#1	Subject#2	Subject#1	Subject#2	Subject#1	Subject#2
2	103	93	107	135	111	140	111	145	112	135	116	125
	15%	20%	15%	15%	15%	15%	15%	15%	15%	15%	15%	15%
Inventory variations w/ Subject#1	$ -	-10	0	28	0	29	0	34	0	23	0	9
Holding inventory	$ -	0	0	28	0	29	0	34	0	23	0	9
Inventory accumulated	$ -	0	0	28	0	57	0	91	0	114	0	123
Holding cost per period	$ -	$ -	$ -	$ 714.00	$ -	$ 739.50	$ -	$ 867.00	$ -	$ 586.50	$ -	$ 229.50
Holding cost accumulated	$ -	$ -	$ -	$ 714.00	$ -	$ 1,453.50	$ -	$ 2,320.50	$ -	$ 2,907.00	$ -	**$ 3,136.50**
Acquisition cost	$ 87,550.00	$ 74,400.00	$ 90,950.00	$ 114,750.00	$ 94,350.00	$ 119,000.00	$ 94,350.00	$ 123,250.00	$ 95,200.00	$ 114,750.00	$ 98,600.00	$ 106,250.00
Acquisition cost (accumulated)	$ 87,550.00	$ 74,400.00	$ 178,500.00	$ 189,150.00	$ 272,850.00	$ 308,150.00	$ 367,200.00	$ 431,400.00	$ 462,400.00	$ 546,150.00	$ 561,000.00	**$ 652,400.00**
Total cost (holding + acquisition)	$ 87,550.00	$ 74,400.00	$ 178,500.00	$ 189,864.00	$ 272,850.00	$ 309,603.50	$ 367,200.00	$ 433,720.50	$ 462,400.00	$ 549,057.00	$ 561,000.00	$ 655,536.50
Shortfalls (unsatisfied demand)	$ -	-10	0	0	0	0	0	0	0	0	0	0
Shortfalls (accumulated)	$ -	-10	0	-10	0	-10	0	-10	0	-10	0	-10
Cost of opportunity (accumul)	$ -	$ (15,500.00)	$ -	$ (15,500.00)	$ -	$ (15,500.00)	$ -	$ (15,500.00)	$ -	$ (15,500.00)	$ -	**$ (15,500.00)**

Observations:

1. Notice that quantities and qualities specified by Subject #1 (Treatment: C1-1) are taken as the reference to compute variations of inventory.
2. For period #1 (Nov-05), variations of quantities specified by subject#2 with respect to those specified by subject#1 are negative, therefore no inventory is carried out for next period (negative variations are discussed in other category of costs).
3. In period #2 (Dec-05), variations are positive (28 pieces). Holding cots for such periods is calculated as follows:

$$I = n \times r \times v \times (1-d) = 28 \times 0.03 \times 1000 \times (1-0.15) = \$714.00$$

The cumulated holding costs for all the periods (planning horizon), and for all products *j* are computed as follows:

$$\sum_{\forall j} \sum_{\forall k} I_k^j$$

For all other treatments (where subject has the extreme expertise – expert decision maker), the computation of holding costs is discussed in subsection 6.2.5.

M5: Acquisition cost

The computation of acquisition costs for the RC-MRP is based on the forecasted quantities and the acquisition discount. The discount depends on the selected raw material supplier. For all the treatments '*acquisition costs*' for any given product *j* are computed as follows:

$$A_k^j = p_k^j \times v^j \times \left(1 - d_k^j\right)$$

Where:

A_k^j = Acquisition cost for period k, for product j

p_k^j = Forecasted quantities of product j, for period k

v^j = Acquisition cost for product j (refer to Appendix G)

d_k^j = Acquisition discount for raw material j during period k,

The cumulated acquisition costs for all the periods (planning horizon), and for all products *j* are computed as follows:

$$\sum_{\forall j} \sum_{\forall k} A_k^j$$

M6: Opportunity cost

The company incurs opportunity costs when it experiences a demand shortfall, i.e., the actual demand (q_k^j) for any given product *j* during period *k* cannot be satisfied neither with the planned quantities (p_k^j) of product (*j*) for the referred period (*k*), nor with the cumulated carrying inventory (n_{k-1}^j) from previous period (k-1). In those cases the company will incur a lost-of-opportunity cost (opportunity cost) for that profit that was not generated for the shortfall.

For all the treatments (except for those where subject has the extreme expertise – expert decision maker), opportunity costs are computed as follows:

$$O_k^j = s_k^j \times v^j \times \left(1 + u^j\right)$$

Where:

O_k^j = Opportunity cost for period k, for product j

s_k^j = Shortfall inventory for period k, for product j

$s_k^j = q_k^j - \left(p_k^j + n_{k-1}^j\right)$; If $q_k^j - \left(p_k^j + n_{k-1}^j\right) \geq 0$;

$s_k^j = 0$; Otherwise

p_k^j = Forecasted quantities of product j, for period k

q_k^j = Actual demand of product j, for period k

n_{k-1}^j = Carrying inventory for period (k-1), for product j

v^j = Acquisition cost for product j (refer to Appendix G)

u^j = Gross margin level for product j (refer to Appendix G)

The cumulated opportunity costs for all the periods (planning horizon), and for all products *j* are computed as follows:

$$\sum_{\forall j} \sum_{\forall k} O_k^j$$

For all other treatments (where subject has the extreme expertise – expert decision maker), the computation of opportunity costs is discussed in subsection 6.2.5.

6.2.5 Correction factors for expert subject

Holding costs and opportunity costs for all subjects with expertise level 2~4 were calculated taking the quantities forecasted by "Expert" decision maker as the reference. This procedure assumes that the output created by "expert" decision maker is correct; consequently, it is assumed that accuracy of expert decision is perfect.

As a consequence of this assumption, output created by "expert" decision maker for the treatments in which he performs incur neither the holding costs (they are zero) nor the opportunity costs.

In order to correct this situation we corrected these values by adding an estimate of the holding and opportunity costs associated with his response. In order to estimate these costs we used the results of the ethnographic study. During the five visits that preceded the experiment, "expert" decision maker solved the RC-MRP problem. During that time the actual demand was unknown. However, at this time the researcher had access to this information. We estimated the holding and opportunity costs associated with the output created by "expert" decision makers during experimental study as follows:

1. Obtain actual demand for the same periods studied during the ethnographic study.

2. Use quantities for all products from actual demand as target quantities to compute inventory deviations on the RC-MRP solutions created by decision maker during ethnographic study.
3. Use inventory deviations from expert's RC-MRP solutions to compute "holding costs" (if excess inventory) or "opportunity costs" (if inventory shortfall). Having actual quantities required by all the periods, and the forecasted production requirements during each period, we proceeded to evaluate each output created by expert. To do this, we took the quantities for actual demand as the real quantities. The quantities forecasted by decision maker were compared against the actual demand. As a result of this, holding and opportunity costs could be computed for each of the outputs created during the ethnographic study. See Table 14 for holding and opportunity costs associated to RC-MRP solutions for each visit.

Table 14: Evaluation of holding and opportunity costs for RC-MRP solutions

Costs / Prod Line	**Actual Demand**					**Planned production**				
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**
Holding cost	0	0	0	0	0	76	50	92	80	92
02" - 06"	0	0	0	0	0	6	4	7	5	8
08" - 12"	0	0	0	0	0	5	5	13	14	8
14" - larger	0	0	0	0	0	65	41	72	61	76
Acquisition cost	27017	31029	29219	24153	23834	27140	30483	29832	23668	24399
02" - 06"	5015	5689	5440	4742	4485	5019	5664	5487	4605	4565
08" - 12"	3509	3944	3281	2654	3016	3143	3806	3237	2645	2834
14" - larger	18493	21396	20498	16757	16333	18978	21013	21108	16418	17000
Opportunity cost	0	0	0	0	0	(2098)	(1342)	(1846)	(1949)	(1589)
02" - 06"	0	0	0	0	0	(77)	(140)	(64)	(303)	(169)
08" - 12"	0	0	0	0	0	(712)	(278)	(423)	(228)	(395)
14" - larger	0	0	0	0	0	(1309)	(924)	(1359)	(1418)	(1025)
Total costs	27,017	31,029	29,219	24,152	23,833	29,310	31,872	31,766	25,693	26,076

Amounts in Thousands of US$

4. The extra holding and opportunity costs computed with the previous procedure are shown in Table 15. Note that no extra acquisition costs are considered.
5. The holding and opportunity costs computed with the previous procedure gave us an idea of how accurate the "expert" decision maker can be for a range of circumstances (recall that circumstances of external environments during each visit of the ethnographic study were different). In order to use these results it was necessary to normalize those values using a useful reference. Since the results for the RC-MRP obtained during experimental study for the expert subject were computed as "Acquisition cost", we decided to normalize extra costs depicted in Table 15 using the acquisition costs as the reference. Table 16 depicts these. In order to use those numbers, we needed to decide how to use those results to correct expert's output during the experimental study.

Table 15: Holding and opportunity costs for RC-MRP solutions

Costs / Prod Line	Extra costs due to variations "Planned" vs "Real"				
	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
Holding cost	76	50	92	80	92
02" - 06"	6	4	7	5	8
08" - 12"	5	5	13	14	8
14" - larger	65	41	72	61	76
Acquisition cost	0	0	0	0	0
02" - 06"					
08" - 12"					
14" - larger					
Opportunity cost	2098	1342	1846	1949	1589
02" - 06"	77	140	64	303	169
08" - 12"	712	278	423	228	395
14" - larger	1309	924	1359	1418	1025
Total costs	2,174	1,392	1,938	2,029	1,681

Amounts in Thousands of US$

Table 16: Extra holding and opportunity costs as % of acquisition cost

	Extra costs (%) using Acquisition (planned) as reference					
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Avg**
% Holding / Total Acquisition Cost (Planned) =	0.28%	0.16%	0.31%	0.34%	0.38%	**0.29%**
% Acquisition / Total Acquisition Cost (Planned) =	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**
% Opportunity / Total Acquisition Cost (Planned) =	7.73%	4.40%	6.19%	8.23%	6.51%	**6.61%**
% Total Extra cost / Total Acquisition Cost (Planned) =	8.01%	4.57%	6.50%	8.57%	6.89%	**6.91%**

6. Reflecting the total % of extra cost into each production line category, we obtained the results depicted in Table 17.
7. Notice the values depicted under the column “Avg” (average) of Table 17. By using these factors we could estimate the expected holding and opportunity costs for the output created by the expert decision maker in the experimental study. These factors are computed as a weighted average of the percentages obtained during the each visit of the ethnographic study.

Table 17: Extra holding and opportunity costs distributed in all production lines

Costs / Prod Line	Extra costs (%) using Acquisition (planned) as reference					
	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Avg**
Holding cost	0.28%	0.16%	0.31%	0.34%	0.38%	**0.29%**
02" - 06"	0.02%	0.01%	0.02%	0.02%	0.03%	**0.02%**
08" - 12"	0.02%	0.02%	0.04%	0.06%	0.03%	**0.03%**
14" - larger	0.24%	0.13%	0.24%	0.26%	0.31%	**0.24%**
Acquisition cost	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**
02" - 06"	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**
08" - 12"	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**
14" - larger	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**
Opportunity cost	7.73%	4.40%	6.19%	8.23%	6.51%	**6.61%**
02" - 06"	0.28%	0.46%	0.21%	1.28%	0.69%	**0.59%**
08" - 12"	2.62%	0.91%	1.42%	0.96%	1.62%	**1.51%**
14" - larger	4.82%	3.03%	4.56%	5.99%	4.20%	**4.52%**
Total costs	8.01%	4.57%	6.50%	8.57%	6.89%	**6.91%**

6.3 Summary

In this chapter we presented the methodology used for the empirical evaluation of the Graph-based Decision Support System proof-of-concept (*miniERP-GDSS*). The evaluation assesses the effectiveness of the Software Assistant in supporting production planners for the task under study, i.e., the creation of a rough-cut material requirements plan (RC-MRP).

Four production planners from the Cooper Cameron Valves manufacturing plant with various levels of expertise in the task under study participated in the evaluation. The production planners participated in two free-time sessions. During one session, participants solved three typical problems for the decision under study utilizing the transactional interface of *miniERP* (control condition). In another session, the participants were asked to solve the set of three problems, but utilizing the *miniERP-GDSS* Decision Support Interface (experimental condition). Participants followed a different order to solve the problems and different sequence of using the interfaces to avoid carry over effects.

A multi-factor analysis of variance model was used to analyze the marginal effects and interaction of the experimental factors. Three factors were studied: (i) interface type, i.e., inclusion or no-inclusion of graph-based model features (*miniERP-GDSS*); (ii) expertise levels of participants; and (iii) difficulty level of exercises. Within each factor, several levels were considered. The high-level of expertise required for the subjects participating in the evaluation imposed constraints on the number of participants available; therefore, only one case per treatment was considered.

The responses of interest (performance measures) were grouped into two categories: (i) benefits to access the required data for the problem at hand, and (ii) benefits to achieve better solutions. For a numeric assessment of solutions, these were compared against historical data.

In Chapter 7 we present an evaluation and discussion of the responses of interest collected during the empirical evaluation.

CHAPTER 7

RESULTS AND DISCUSSION

Descriptive and detailed analysis of results collected during the experimental study is presented in this chapter. Descriptive analysis provides an overview of main findings of the empirical evaluation. A more detailed factor analysis includes the evaluation of treatment means tables and plots. Finally, a description of the findings completes this chapter.

7.1 Overview of results

In this section we provide an overview of the results obtained during the empirical evaluation. Results for five experimental responses classified into two types are presented: (1) *transactional results* that reflect the interaction of the subjects with the software implementation for all the treatments considered in the experimental study; and (2) *numerical results* that reflect the quantitative accuracy of the solution for the problem at hand, e.g., accuracy to fulfill problem's goals such as minimum costs.

7.1.1 Overview of transactional results

Transactional results are divided into two groups. Results reflecting transactional features of interaction with control and experimental conditions are depicted in Table 18 and Table 19, respectively.

From Table 18 it can be seen that subject #1 (expertise level 'Extreme') performed faster (shorter execution times, i.e., M1.1) when using the experimental condition, i.e., when using the implementation software. For instance when 'Problem_Conds' are favorable execution times decreased from 150.80 min to 111.66

minutes. Similar improvements from control to experimental conditions can be observed for other performance measures: ‘Search for documents’ (M1.2) and the ratio of ‘Search for documents over the total execution time’ (M1.4). Similar observations can be drawn for subjects #2 ~ #4 (non-experts) (see subject #2 in Table 18 and subjects #3 and #4 in Table 19).

In general, it is noticeable that the decision aid provided by the experimental system (*miniERP-GDSS*) improved transactional results for all subjects, i.e., they accessed the needed data in shorter times and with a significant reduction of data overload.

7.1.2 <u>Overview of numeric results</u>

Summary of numerical results for control (*miniERP*) and experimental (*miniERP-GDSS*) conditions are depicted in Table 20 and Table 21, respectively.

From Table 20, it can be seen that subject #1 (expertise level ‘Extreme’) showed improvement in both cost-based performance measures (‘Total cost’ and ‘Percentage of Improvement’) when using the experimental condition (except for the case when “Problem_Conds’ are favorable). For instance when ‘Problem_Conds’ are ‘Adverse’ total costs decreased from $3141 to $3022 (-4.1%).

Subjects #2 ~ #4 (non-experts) (see subject #2 in Table 20 and subjects #3 and #4 in Table 21) showed small improvements in both cost-based performance measures for all values of ‘Problem Condition’ factor. The causes for these improvements are not clear because different paradigms for solving the problems affected the quantitative results. Further possible explanations for these variations are provided later in this chapter.

Table 18: Transactional results – first part: subjects **#1** and **#2**

Experimental Factor												
#3: Expertise (Subjects' expertise level)	Subject #1 (Extreme expertise)						Subject #2 (High expertise)					
#2: Condition (Problem's condition)	Favorable		Adverse		Mixed		Favorable		Adverse		Mixed	
#1: Interface (Interface type) (1)	C1 (1)	E1	C2	E2	C3	E3	C1	E1	C2	E2	C3	E3
M1. Execution time [minutes]:												
M1.1 Complete solution M1.1 = M1.2 + M1.3	150.80	111.66	152.76	114.76	136.68	101.22	110.38	81.38	108.10	89.10	91.10	76.10
M1.2 Creating/browsing documents	44.66	18.29	69.43	24.43	52.62	20.62	82.92	27.92	72.28	28.28	62.91	21.91
M1.3 Using relevant documents	106.13	93.37	83.33	90.33	84.07	80.60	27.47	53.47	35.82	60.82	28.19	54.19
M1.4 Ratio: (M1.2) / (M1.1) (2)	30%	16%	45%	21%	38%	20%	75%	34%	67%	32%	69%	29%
M2. Data overload level:												
M2.1 No.Documents **created** (C) or **browsed** (E)	72	43	60	54	51	48	85	39	66	48	61	43
M2.2 No.Documents **used**	42	40	48	51	43	44	35	32	43	38	37	39
M2.3 Ratio: (M2.1) / (M2.2) (3)	1.71	1.08	1.25	1.06	1.19	1.09	2.43	1.22	1.53	1.26	1.65	1.10
M3. Completeness of Expert Knowledge database (EK):												
M3.1 No.Documents **created** & **used** in (C), but not available for **browsing** in (E), missing in (EK)	0	n/a	3	n/a	3	n/a	6	n/a	8	n/a	7	n/a
M3.2 Ratio: (M3.1) / (M2.2) (5)	0%	n/a	6%	n/a	7%	n/a	17%	n/a	19%	n/a	19%	n/a

Notes:

(1) A three-factor ANOVA model is proposed for the analysis. Three cases under two conditions are being tested: Control Condition (C) and Experimental Condition (E). Control Condition (C) - uses the *miniERP* interface; Experimental Condition (E) - uses the *miniERP-GDSS* interface.

(2) The (M1.2)/(M1.1) ratio represents the percentage of time spent creating/browsing documents

(3) The (M2.1)/(M2.2) ratio reveals the amount of unnecessary documents created/browsed during the search of data. A ratio closer to 1 indicates that most of documents created/browsed resulted useful for the decision.

(4) The M3.1 variable reveals the completeness of the knowledge base (the core of the *miniERP-GDSS* interface). Ideally, a value of 'zero' guarantees total completeness, i.e., no useful document was left out of EK.

(5) The (M3.1)/(M2.2) ratio presents the completeness of knowledge base. Same as (M3.1), but as a % of M2.2 (the total number of documents used to compute a solution)

Table 19: Transactional results – second part: subjects **#3** and **#4**

Experimental Factor												
#3: Expertise (Subjects' expertise level)	Subject #3 (Medium expertise)						Subject #4 (Low expertise)					
#2: Condition (Problem's condition)	Favorable		Adverse		Mixed		Favorable		Adverse		Mixed	
#1: Interface (Interface type) ([1])	C1	E1	C2	E2	C3	E3	C1	E1	C2	E2	C3	E3
M1. Execution time [minutes]:												
M1.1 Complete solution M1.1 = M1.2 + M1.3	121.03	106.03	131.27	116.27	127.22	89.77	131.37	116.37	151.27	136.27	125.05	110.05
M1.2 Creating/browsing documents	57.65	28.65	82.85	23.85	78.70	5.67	52.65	37.65	70.85	55.85	51.42	36.42
M1.3 Using relevant documents	63.38	77.38	48.42	92.42	48.52	84.10	78.72	78.72	80.42	80.42	73.63	73.63
M1.4 Ratio: (M1.2) / (M1.1) ([2])	48%	27%	63%	21%	62%	6%	40%	32%	47%	41%	41%	33%
M2. Data overload level:												
M2.1 No.Documents **created** (C) or **browsed** (E)	57	35	74	44	58	36	66	36	48	26	39	19
M2.2 No.Documents **used**	25	27	41	24	32	26	30	19	34	16	32	16
M2.3 Ratio: (M2.1) / (M2.2) ([3])	2.28	1.30	1.80	1.83	1.81	1.38	2.20	1.89	1.41	1.63	1.22	1.19
M3. Completeness of Expert Knowledge database (EK):												
M3.1 No.Documents **created** & **used** in (C), but not available for **browsing** in (E), missing in (EK)	1	n/a	2	n/a	2	n/a	0	n/a	1	n/a	0	n/a
M3.2 Ratio: (M3.1) / (M2.2) ([5])	4%	n/a	5%	n/a	6%	n/a	0%	n/a	3%	n/a	0%	n/a

Notes:

(1) A three-factor ANOVA model is proposed for the analysis. Three cases under two conditions are being tested: Control Condition (C) and Experimental Condition (E). Control Condition (C) - uses the *miniERP* interface; Experimental Condition (E) - uses the *miniERP-GDSS* interface.

(2) The (M1.2)/(M1.1) ratio represents the percentage of time spent creating/browsing documents

(3) The (M2.1)/(M2.2) ratio reveals the amount of unnecessary documents created/browsed during the search of data. A ratio closer to 1 indicates that most of documents created/browsed resulted useful for the decision.

(4) The M3.1 variable reveals the completeness of the knowledge base (the core of the *miniERP-GDSS* interface). Ideally, a value of 'zero' guarantees total completeness, i.e., no useful document was left out of EK.

(5) The (M3.1)/(M2.2) ratio presents the completeness of knowledge base. Same as (M3.1), but as a % of M2.2 (the total number of documents used to compute a solution)

Table 20: Numerical results – first part: subjects **#1** and **#2**

Costs	Actual Demand			Subject #1 (Extreme expertise)						Subject #2 (High expertise)					
	Favorable	Adverse	Mixed	Favorable		Adverse		Mixed		Favorable		Adverse		Mixed	
				C1	E1	C2	E2	C3	E3	C1	E1	C2	E2	C3	E3
Holding cost				**63**	**63**	**9**	**9**	**54**	**51**	**78**	**105**	**12**	**8**	**76**	**74**
02" - 06"				5	5	1	1	5	4	12	7	3	2	29	29
08" - 12"				8	8	2	1	7	6	7	5	3	2	19	18
14" - larger				51	51	7	7	43	42	60	94	7	5	30	27
Acquisition cost	**21195**	**2882**	**17724**	**21135**	**21256**	**2938**	**2827**	**18089**	**17361**	**20764**	**21149**	**2507**	**1980**	**17915**	**16413**
02" - 06"	5261	743	3923	5250	5272	708	780	3993	3853	4031	5005	670	689	3754	3491
08" - 12"	2931	636	2521	2890	2973	728	545	2496	2546	2782	2586	561	262	2447	2496
14" - larger	13003	1503	11281	12996	13011	1503	1503	11600	10962	13952	13558	1277	1030	11716	10427
Opportunity cost				**(1398)**	**(1406)**	**(195)**	**(187)**	**(1197)**	**(1149)**	**(2226)**	**(608)**	**(1981)**	**(2022)**	**(1424)**	**(1474)**
02" - 06"				(124)	(125)	(18)	(17)	(107)	(102)	(1640)	(428)	(298)	(383)	(719)	(744)
08" - 12"				(319)	(321)	(45)	(43)	(273)	(262)	(300)	(180)	(679)	(634)	(125)	(130)
14" - larger				(956)	(961)	(133)	(128)	(818)	(785)	(286)	0	(1006)	(1006)	(580)	(601)
Total costs	**21195**	**2882**	**17724**	**22594**	**22724**	**3141**	**3022**	**19338**	**18560**	**23067**	**21860**	**4500**	**4010**	**19414**	**17959**

% improv C vs E		1%		-4%		-4%		-5%		-11%		-7%
% Improv vs Actual	6.6%	7.2%	9.0%	4.9%	9.1%	4.7%	8.8%	3.1%	56.1%	39.1%	9.5%	1.3%

Amounts in <u>Thousands</u> of US$

Table 21: Numerical results – second part: subjects **#3** and **#4**

Costs	Actual Demand			Subject #3 (Medium expertise)						Subject #4 (Low expertise)					
	Favorable	Adverse	Mixed	Favorable		Adverse		Mixed		Favorable		Adverse		Mixed	
				C1	E1	C2	E2	C3	E3	C1	E1	C2	E2	C3	E3
Holding cost				**204**	**112**	**3**	**2**	**258**	**267**	**129**	**73**	**77**	**76**	**229**	**237**
02" - 06"				82	37	2	2	40	41	8	5	22	20	72	75
08" - 12"				109	58	1	1	151	156	7	6	20	21	9	10
14" - larger				14	17	0	0	68	70	115	63	36	36	148	154
Acquisition cost	**21195**	**2882**	**17724**	**23598**	**23637**	**1747**	**1578**	**19029**	**17648**	**24313**	**21586**	**5406**	**5002**	**18490**	**16799**
02" - 06"	5261	743	3923	7625	6192	602	671	4113	3969	4962	4269	1534	1534	4113	3969
08" - 12"	2931	636	2521	6501	4890	482	244	2621	2490	2944	2924	1413	1196	2546	2419
14" - larger	13003	1503	11281	9473	12556	663	663	12296	11189	16408	14395	2460	2274	11832	10412
Opportunity cost				**(7491)**	**(2381)**	**(3462)**	**(3595)**	**(2190)**	**(2266)**	**(541)**	**(1515)**	**(1099)**	**(1024)**	**(2103)**	**(2176)**
02" - 06"				(350)	(308)	(489)	(521)	(400)	(414)	(422)	(1090)	(164)	(179)	(1398)	(1447)
08" - 12"				0	0	(778)	(878)	(50)	(52)	(119)	(141)	(90)	0	(125)	(130)
14" - larger				(7141)	(2074)	(2196)	(2196)	(1740)	(1801)	0	(286)	(846)	(846)	(580)	(601)
Total costs	**21195**	**2882**	**17724**	**31291**	**26128**	**5210**	**5174**	**21476**	**20180**	**24982**	**23174**	**6581**	**6102**	**20821**	**19212**

% improv C vs E		-16%		-1%		-6%		-7%		-7%		-8%
% Improv vs Actual	47.6%	23.3%	80.8%	79.5%	21.2%	13.9%	17.9%	9.3%	128.4%	111.7%	17.5%	8.4%

Amounts in <u>Thousands</u> of US$

7.2 Analysis of results

In the following subsections we present a more detailed analysis of results. A multi-factor Analysis of Variance Model (ANOVA) was used to determine marginal and interaction effects of experimental factors. Treatment means tables for responses under study were built. Treatment means plots are provided to examine main effects and interaction effects. In order to support the subjective assessment of main effects and interaction effects we use the sum of squares concept (SS) and ratio of sum of squares (%SS).

The sum of squares is a useful concept to provide a measure of the deviation of observations from treatment means. To derive this, we start from the total deviation of an observation (a response from a treatment) Y_{ijk} from the overall mean $\mu_{...}$ in two stages. First, we obtain a decomposition of the total deviation $Y_{ijk} - \mu_{...}$ by viewing the study as consisting of *ab* treatments:

(7.1) $Y_{ijk} - \mu_{...} = \mu_{ij.} - \mu_{...} + Y_{ijk} - \mu_{ij.}$

When we square (7.1) and sum over all cases, the cross product term drops out and we obtain:

(7.2) $SSTO = SSTR + SSE$

SSTR reflects the variability between the *ab* estimated treatment means and is the ordinary *treatment sum of squares*, and SSE reflects the variability within treatments and is the usual *error sum of squares*. The three subscripts are used to designate a treatment (one for each experimental factor). If we decompose the estimated treatment mean deviation $\mu_{ij.} - \mu_{...}$ in terms of components reflecting the factors A main effect, B main effect, C main effect, and the corresponding interactions, that is, AB, AC, BC and ABC we obtain:

(7.3) $SSTO = SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC + SSE$

SSA, SSB, and SSC are the usual main effects sums of squares. For instance, the larger (absolutely) are the estimated main B effects ($\mu_{.j.} - \mu_{...}$), the larger will be SSB. SSAB, SSAC, and SSBC are the usual two-factor interactions sum of squares. For instance, the larger (absolutely) are the estimated AB interactions ($\mu_{ij.} - \mu_{i..} - \mu_{.j.} + \mu_{...}$), the larger will be SSAB. Finally, SSABC is the three-factor interactions sum of squares. The larger (absolutely) are these estimated three-factor interactions, the larger will be SSABC, and consequently so will be the SSABC ratio.

The assumption of normal distribution could not be satisfied; therefore we did not perform hypotheses testing or F tests.

7.2.1 Analysis of factor effects for "Total Execution Times"

The "Total Execution Times" response refers to the total time subjects needed to complete each RC-MRP exercise. The Treatment Means Table for this response is provided in Table 22. Using this table, we built Treatment Means Plots for main effects and Interaction effects; these are presented in Figure 28 and Figure 29, respectively. The Analysis of Variance (ANOVA) table for this response is presented in Figure 30. A subjective examination of the marginal and interaction effects using these plots and the ANOVA table complete this subsection.

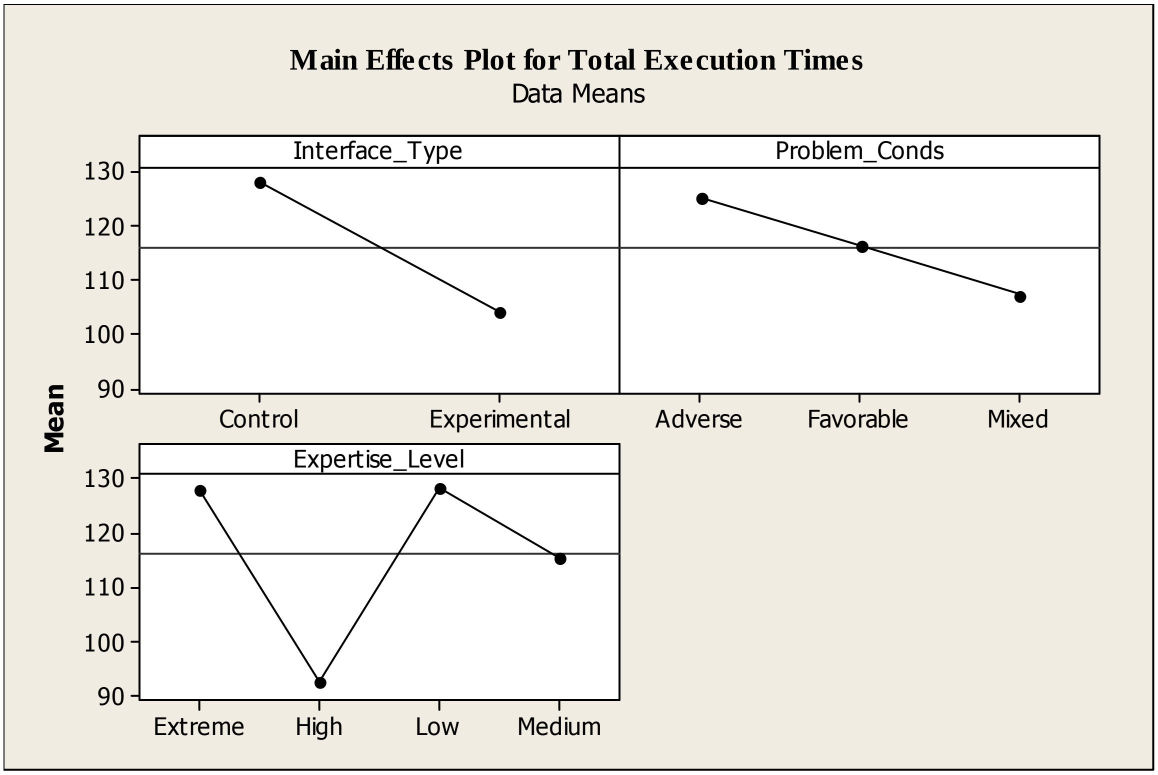


Figure 28: Main effects plot for "Total Execution Times"

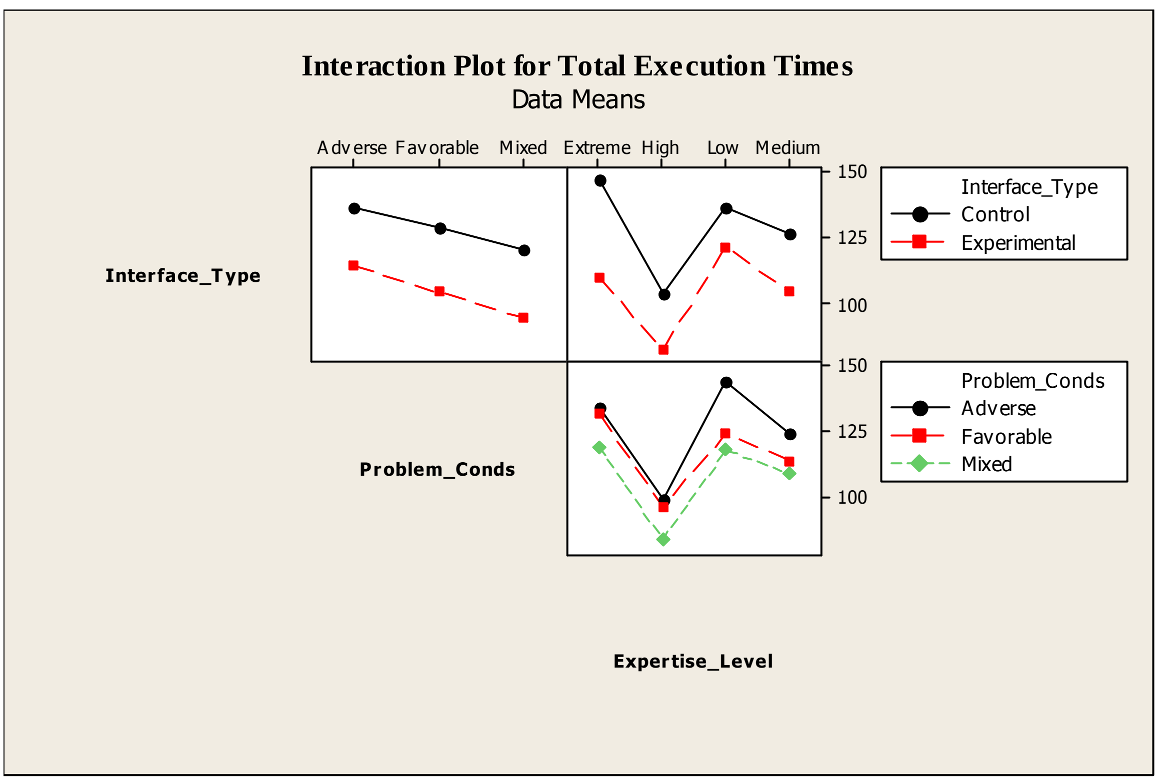


Figure 29: Interaction effects plot for "Total Execution Times"

Table 22: Transactional results – Treatment Means Tables

Mean *Total Execution Time* according to Interface type, Problems' conditions, and Expertise level

Interface type (Factor A)	Subjets' Expertise Level (Factor C) and Problem's Conditions (Factor B)																			
	k=1; Extreme				k=2; High				k=3; Medium				k=4; Low				Average			
	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average
i =1; Control(C)	150.80 (μ_{111})	152.76 (μ_{121})	136.68 (μ_{131})	146.75 $(\mu_{1\cdot 1})$	110.38 (μ_{112})	108.10 (μ_{122})	91.10 (μ_{132})	103.19 $(\mu_{1\cdot 2})$	121.03 (μ_{113})	131.27 (μ_{123})	127.22 (μ_{133})	126.51 $(\mu_{1\cdot 3})$	131.37 (μ_{114})	151.27 (μ_{124})	125.05 (μ_{134})	135.89 $(\mu_{1\cdot 4})$	128.39 $(\mu_{11\cdot})$	135.85 $(\mu_{12\cdot})$	120.01 $(\mu_{13\cdot})$	128.08 $(\mu_{1\cdot\cdot})$
i =2; Experimental (E)	111.66 (μ_{211})	114.76 (μ_{221})	101.22 (μ_{231})	109.21 $(\mu_{2\cdot 1})$	81.38 (μ_{212})	89.10 (μ_{222})	76.10 (μ_{232})	82.19 $(\mu_{2\cdot 2})$	106.03 (μ_{213})	116.27 (μ_{223})	89.77 (μ_{233})	104.02 $(\mu_{2\cdot 3})$	116.37 (μ_{214})	136.27 (μ_{224})	110.05 (μ_{234})	120.89 $(\mu_{2\cdot 4})$	103.86 $(\mu_{21\cdot})$	114.10 $(\mu_{22\cdot})$	94.28 $(\mu_{23\cdot})$	104.08 $(\mu_{2\cdot\cdot})$
Average	131.23 $(\mu_{\cdot 11})$	133.76 $(\mu_{\cdot 21})$	118.95 $(\mu_{\cdot 31})$	127.98 $(\mu_{\cdot\cdot 1})$	95.88 $(\mu_{\cdot 12})$	98.60 $(\mu_{\cdot 22})$	83.60 $(\mu_{\cdot 32})$	92.69 $(\mu_{\cdot\cdot 2})$	113.53 $(\mu_{\cdot 13})$	123.77 $(\mu_{\cdot 23})$	108.49 $(\mu_{\cdot 33})$	115.26 $(\mu_{\cdot\cdot 3})$	123.87 $(\mu_{\cdot 14})$	143.77 $(\mu_{\cdot 24})$	117.55 $(\mu_{\cdot 34})$	128.39 $(\mu_{\cdot\cdot 4})$	116.13 $(\mu_{\cdot 1\cdot})$	124.97 $(\mu_{\cdot 2\cdot})$	107.15 $(\mu_{\cdot 3\cdot})$	116.08 $(\mu_{\cdot\cdot\cdot})$

Time units [minutes]

Mean *Search for Documents Time* according to Interface type, Problems' conditions, and Expertise level

Interface type (Factor A)	Subjets' Expertise Level (Factor C) and Problem's Conditions (Factor B)																			
	k=1; Extreme				k=2; High				k=3; Medium				k=4; Low				Average			
	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average
i =1; Control(C)	44.66 (μ_{111})	69.43 (μ_{121})	52.62 (μ_{131})	55.57 $(\mu_{1\cdot 1})$	82.92 (μ_{112})	72.28 (μ_{122})	62.91 (μ_{132})	72.70 $(\mu_{1\cdot 2})$	57.65 (μ_{113})	82.85 (μ_{123})	78.70 (μ_{133})	73.07 $(\mu_{1\cdot 3})$	52.65 (μ_{114})	70.85 (μ_{124})	51.42 (μ_{134})	58.31 $(\mu_{1\cdot 4})$	59.47 $(\mu_{11\cdot})$	73.85 $(\mu_{12\cdot})$	61.41 $(\mu_{13\cdot})$	64.91 $(\mu_{1\cdot\cdot})$
i =2; Experimental (E)	18.29 (μ_{211})	24.43 (μ_{221})	20.62 (μ_{231})	21.11 $(\mu_{2\cdot 1})$	27.92 (μ_{212})	28.28 (μ_{222})	21.91 (μ_{232})	26.04 $(\mu_{2\cdot 2})$	28.65 (μ_{213})	23.85 (μ_{223})	5.67 (μ_{233})	19.39 $(\mu_{2\cdot 3})$	37.65 (μ_{214})	55.85 (μ_{224})	36.42 (μ_{234})	43.31 $(\mu_{2\cdot 4})$	28.13 $(\mu_{21\cdot})$	33.10 $(\mu_{22\cdot})$	21.15 $(\mu_{23\cdot})$	27.46 $(\mu_{2\cdot\cdot})$
Average	31.47 $(\mu_{\cdot 11})$	46.93 $(\mu_{\cdot 21})$	36.62 $(\mu_{\cdot 31})$	38.34 $(\mu_{\cdot\cdot 1})$	55.42 $(\mu_{\cdot 12})$	50.28 $(\mu_{\cdot 22})$	42.41 $(\mu_{\cdot 32})$	49.37 $(\mu_{\cdot\cdot 2})$	43.15 $(\mu_{\cdot 13})$	53.35 $(\mu_{\cdot 23})$	42.18 $(\mu_{\cdot 33})$	46.23 $(\mu_{\cdot\cdot 3})$	45.15 $(\mu_{\cdot 14})$	63.35 $(\mu_{\cdot 24})$	43.92 $(\mu_{\cdot 34})$	50.81 $(\mu_{\cdot\cdot 4})$	43.80 $(\mu_{\cdot 1\cdot})$	53.48 $(\mu_{\cdot 2\cdot})$	41.28 $(\mu_{\cdot 3\cdot})$	46.19 $(\mu_{\cdot\cdot\cdot})$

Time units [minutes]

Mean *Ratio: (Search for Documents) / (Total Execution) Times* according to Interface type, Problems' conditions, and Expertise level

Interface type (Factor A)	Subjets' Expertise Level (Factor C) and Problem's Conditions (Factor B)																			
	k=1; Extreme				k=2; High				k=3; Medium				k=4; Low				Average			
	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average
i =1; Control(C)	30% (μ_{111})	45% (μ_{121})	38% (μ_{131})	37.86% $(\mu_{1\cdot 1})$	75% (μ_{112})	67% (μ_{122})	69% (μ_{132})	70.35% $(\mu_{1\cdot 2})$	48% (μ_{113})	63% (μ_{123})	62% (μ_{133})	57.54% $(\mu_{1\cdot 3})$	40% (μ_{114})	47% (μ_{124})	41% (μ_{134})	42.68% $(\mu_{1\cdot 4})$	0.48 $(\mu_{11\cdot})$	0.56 $(\mu_{12\cdot})$	0.53 $(\mu_{13\cdot})$	52.10% $(\mu_{1\cdot\cdot})$
i =2; Experimental (E)	16% (μ_{211})	21% (μ_{221})	20% (μ_{231})	19.34% $(\mu_{2\cdot 1})$	34% (μ_{212})	32% (μ_{222})	29% (μ_{232})	31.61% $(\mu_{2\cdot 2})$	27% (μ_{213})	21% (μ_{223})	6% (μ_{233})	17.95% $(\mu_{2\cdot 3})$	32% (μ_{214})	41% (μ_{224})	33% (μ_{234})	35.48% $(\mu_{2\cdot 4})$	0.28 $(\mu_{21\cdot})$	0.29 $(\mu_{22\cdot})$	0.22 $(\mu_{23\cdot})$	26.10% $(\mu_{2\cdot\cdot})$
Average	23.00% $(\mu_{\cdot 11})$	33.37% $(\mu_{\cdot 21})$	29.43% $(\mu_{\cdot 31})$	28.60% $(\mu_{\cdot\cdot 1})$	54.71% $(\mu_{\cdot 12})$	49.31% $(\mu_{\cdot 22})$	48.93% $(\mu_{\cdot 32})$	50.98% $(\mu_{\cdot\cdot 2})$	37.33% $(\mu_{\cdot 13})$	41.81% $(\mu_{\cdot 23})$	34.09% $(\mu_{\cdot 33})$	37.74% $(\mu_{\cdot\cdot 3})$	36.22% $(\mu_{\cdot 14})$	43.91% $(\mu_{\cdot 24})$	37.10% $(\mu_{\cdot 34})$	39.08% $(\mu_{\cdot\cdot 4})$	37.81% $(\mu_{\cdot 1\cdot})$	42.10% $(\mu_{\cdot 2\cdot})$	37.39% $(\mu_{\cdot 3\cdot})$	39.10% $(\mu_{\cdot\cdot\cdot})$

Mean *Ratio: (#Docs Used) / (#Docs Created/Browsed)* according to Interface type, Problems' conditions, and Expertise level

Interface type (Factor A)	Subjets' Expertise Level (Factor C) and Problem's Conditions (Factor B)																			
	k=1; Extreme				k=2; High				k=3; Medium				k=4; Low				Average			
	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average
i =1; Control(C)	1.71 (μ_{111})	1.25 (μ_{121})	1.19 (μ_{131})	1.38 $(\mu_{1\cdot 1})$	2.43 (μ_{112})	1.53 (μ_{122})	1.65 (μ_{132})	1.87 $(\mu_{1\cdot 2})$	2.28 (μ_{113})	1.80 (μ_{123})	1.81 (μ_{133})	1.97 $(\mu_{1\cdot 3})$	2.20 (μ_{114})	1.41 (μ_{124})	1.22 (μ_{134})	1.61 $(\mu_{1\cdot 4})$	2.16 $(\mu_{11\cdot})$	1.50 $(\mu_{12\cdot})$	1.47 $(\mu_{13\cdot})$	1.71 $(\mu_{1\cdot\cdot})$
i =2; Experimental (E)	1.08 (μ_{211})	1.06 (μ_{221})	1.09 (μ_{231})	1.07 $(\mu_{2\cdot 1})$	1.22 (μ_{212})	1.26 (μ_{222})	1.10 (μ_{232})	1.19 $(\mu_{2\cdot 2})$	1.30 (μ_{213})	1.83 (μ_{223})	1.38 (μ_{233})	1.50 $(\mu_{2\cdot 3})$	1.89 (μ_{214})	1.63 (μ_{224})	1.19 (μ_{234})	1.57 $(\mu_{2\cdot 4})$	1.37 $(\mu_{21\cdot})$	1.45 $(\mu_{22\cdot})$	1.19 $(\mu_{23\cdot})$	1.34 $(\mu_{2\cdot\cdot})$
Average	1.39 $(\mu_{\cdot 11})$	1.15 $(\mu_{\cdot 21})$	1.14 $(\mu_{\cdot 31})$	1.23 $(\mu_{\cdot\cdot 1})$	1.82 $(\mu_{\cdot 12})$	1.40 $(\mu_{\cdot 22})$	1.38 $(\mu_{\cdot 32})$	1.53 $(\mu_{\cdot\cdot 2})$	1.79 $(\mu_{\cdot 13})$	1.82 $(\mu_{\cdot 23})$	1.60 $(\mu_{\cdot 33})$	1.74 $(\mu_{\cdot\cdot 3})$	2.05 $(\mu_{\cdot 14})$	1.52 $(\mu_{\cdot 24})$	1.20 $(\mu_{\cdot 34})$	1.59 $(\mu_{\cdot\cdot 4})$	1.76 $(\mu_{\cdot 1\cdot})$	1.47 $(\mu_{\cdot 2\cdot})$	1.33 $(\mu_{\cdot 3\cdot})$	1.52 $(\mu_{\cdot\cdot\cdot})$

```
Response: Total Execution Times
Interface_Type, Problem_Conditions, Expertise_Level

Factor           Type   Levels  Values
Interface_Type   fixed       2  Control, Experimental
Problem_Conds    fixed       3  Adverse, Favorable, Mixed
Expertise_Level  fixed       4  Extreme, High, Low, Medium


Analysis of Variance for Exec_Times, using Adjusted SS for Tests

Source                          DF    Seq SS   Adj SS   Adj MS   F  P
Interface_Type                   1   3457.24  3457.24  3457.24  **
Problem_Conds                    2   1270.92  1270.92   635.46  **
Expertise_Level                  3   5044.60  5044.60  1681.53  **
Interface_Type*Problem_Conds     2     16.67    16.67     8.34  **
Interface_Type*Expertise_Level   3    413.20   413.20   137.73  **
Problem_Conds*Expertise_Level    6    226.65   226.65    37.77  **
Interface_Type*Problem_Conds*    6    206.85   206.85    34.48  **
  Expertise_Level
Error                            0         *        *        *
Total                           23  10636.13
```

Figure 30: ANOVA table for “Total Execution Times”

7.2.1.1 Recognition of main effects

For a first recognition of main effects, we can use treatment means curves in Figure 28; all these curves have a non-zero slope, which indicates the presence of main effects of all the experimental factors. The sum of squares concept provides a quantitative insight of the degree of importance of each marginal effect. From Figure 30 we can see that the two most significant factors for the response are the ‘Expertise level’ factor (SS = 5044.60) and the ‘Interface Type’ factor (SS = 3457.24). The ‘Problem Conditions’ factor has relatively a minor importance (SS = 1270.92), which can be also noticed from the treatment means curve since the difference in height for different values of the factor is not so big. For a percentage comparison of the sum of squares concept refer to the discussion of results presented in the next section (Section 7.3).

7.2.1.2 Recognition of interaction effects

A visual inspection of whether the treatment means curves for the different factor levels in a treatment means plot are parallel is helpful to recognize the presence of interactions. Figure 29 depicts a plot of the interaction plot for Total Execution Times. Note that the treatment means curves for the three factors are not completely parallel, which indicates some degree of interactions.

Although a visual inspection reflects the presence of interaction effects, the use of the Sum of Squares concept provides a better insight into the magnitude and importance of each interaction. From Figure 30 we can see that although interactions are smaller than marginal effects, they are also present. It is interesting to note the interaction between 'Interface type' and 'Expertise level' (SS = 413.20) (notice that the most important marginal effects come from these two factors). If we observe the treatment means curves for the interaction of these factors we notice that for the expertise level = "Extreme', the effect of the interface type is the largest. This makes perfect sense if we recall that the knowledge base underlying the software implementation was built with the input of the subject with extreme expertise level. For the other treatments (other subjects) we can also observe significant impact of the interface type.

7.2.2 <u>Analysis of factor effects for "Documents Search Times"</u>

The "Documents Search Times" response refers to the time subjects needed to search for documents to create a solution for each RC-MRP exercise. Table 22 (part B) displays the Treatment Means Table for this response. Treatment Means Plots for main effects and Interaction effects are presented in Figure 31 and Figure 32, respectively. The Analysis of Variance (ANOVA) table for this response is presented in Figure 33. These figures and a subjective examination of the marginal and interaction effects are presented next.

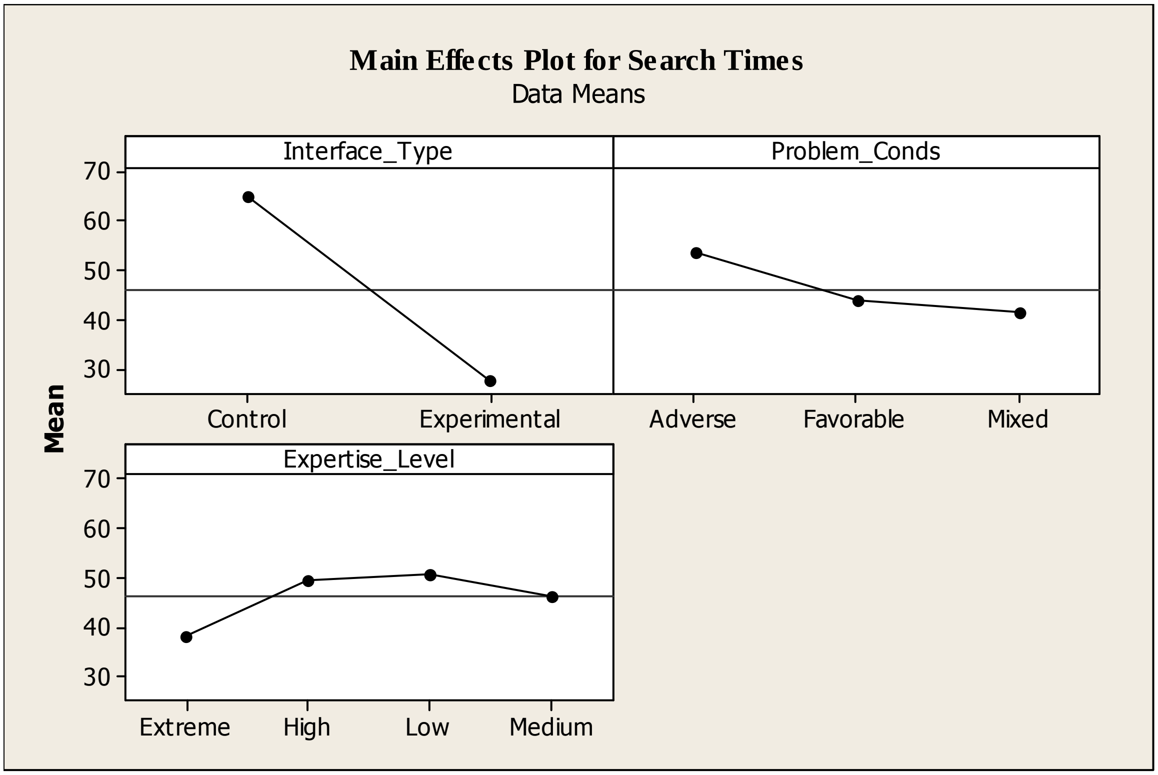


Figure 31: Main effects plot for “Documents Search Times”

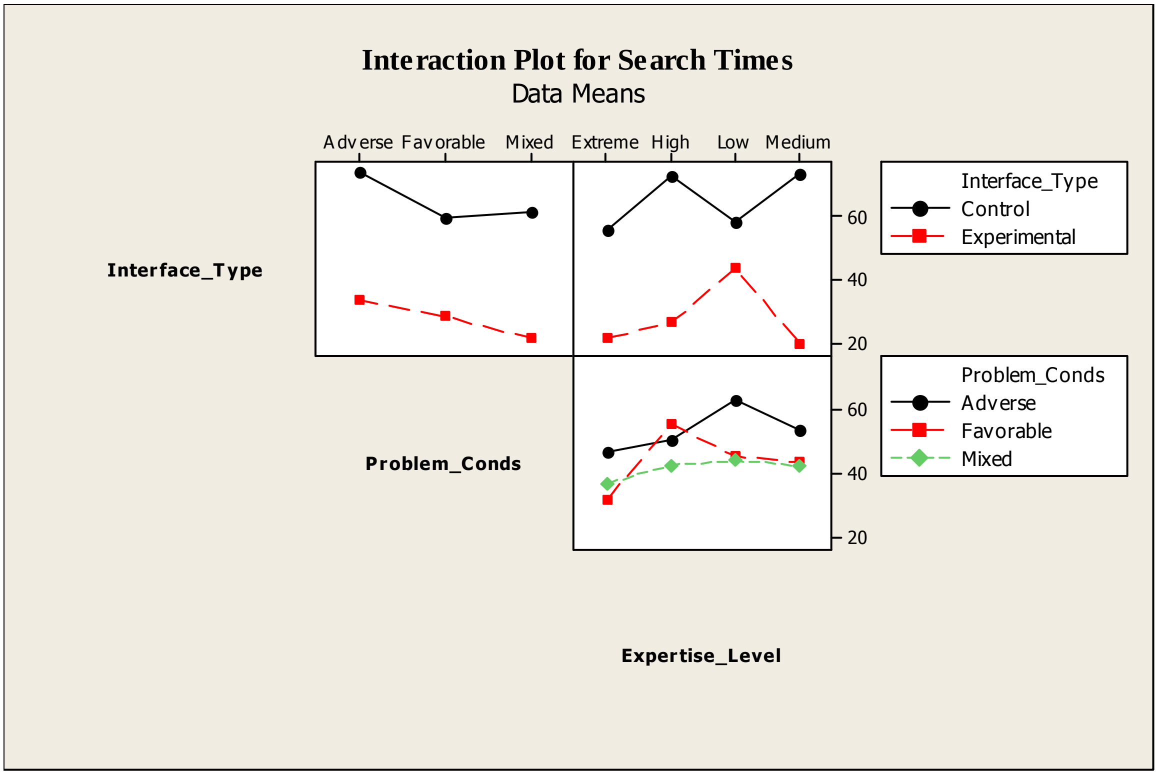


Figure 32: Interaction effects plot for "Documents Search Times"

Documents Search Times

Interface_Type, Problem_Conditions, Expertise_Level

Factor	Type	Levels	Values
Interface_Type	fixed	2	Control, Experimental
Problem_Conds	fixed	3	Adverse, Favorable, Mixed
Expertise_Level	fixed	4	Extreme, High, Low, Medium

Analysis of Variance for Search_time, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface_Type	1	8415.51	8415.51	8415.51	**	
Problem_Conds	2	663.37	663.37	331.69	**	
Expertise_Level	3	558.25	558.25	186.08	**	
Interface_Type*Problem_Conds	2	112.10	112.10	56.05	**	
Interface_Type*Expertise_Level	3	1291.85	1291.85	430.62	**	
Problem_Conds*Expertise_Level	6	382.66	382.66	63.78	**	
Interface_Type*Problem_Conds* Expertise_Level	6	539.43	539.43	89.91	**	
Error	0	*	*	*		
Total	23	11963.17				

Figure 33: ANOVA table for "Documents Search Times"

7.2.2.1 Recognition of main effects

From Figure 31 we observe that each of the treatment means curves has a non-zero slope, which indicates the presence of marginal effects of all the experimental factors. With a visual inspection it is noticeable that the treatment means curve for the ‘Interface Type’ factor presents the largest differential in heights. This fact is confirmed using the sum of squares, which provides a quantitative insight of the degree of importance of each marginal effect. From Figure 33 we can see that the ‘Interface Type’ constitutes the most significant factor (SS = 8415.51). For a percentage comparison of the sum of squares, refer to the discussion of results presented in the next section (Section 7.3).

7.2.2.2 Recognition of interaction effects

Again, from a visual inspection we can see a lack of parallelism between the treatments means curves for the different factor levels; this fact indicates the presence of interactions. Figure 32 presents the interaction plot for Document Search Times. Note that the treatment means curves for the three factors are not parallel.

Although a visual inspection reflects the presence of interaction effects, the use of the Sum of Squares concept provides a better insight into the magnitude and importance of each interaction. From Figure 33 we can see that although interactions are smaller than marginal effects, there are also present. We can observe an important interaction between the ‘Interface type’ and ‘Expertise level’ factors. In this case, sum of squares is larger than the other interactions (SS = 1291.85, %SS = 10.79%). For a percentage comparison of the sum of squares concept refer to the discussion of results presented in the next section (Section 7.3).

7.2.3 Analysis of factor effects for "Ratio: Search / Total (Times)"

The Ratio of the times needed for "Documents Search over Total Execution" is a response that provides a normalization of the previous response (Document Search Times). This ratio refers to the time subjects needed to search for documents to create a solution for each RC-MRP exercise, but as a difference from the previous response, this factor considers the skill level of the user. For this reason, it is expected that this response (Ratio: Search/Total Execution) will better reflect the effects of experimental factors. The Treatment Means Table for this response is provided in Table 22. Using this table, we built Treatment Means Plots for main effects and Interaction effects; these are presented in Figure 34 and Figure 35, respectively. The Analysis of Variance (ANOVA) table for this response is presented in Figure 36. A subjective examination of the marginal and interaction effects using these plots and the ANOVA table completes this subsection.

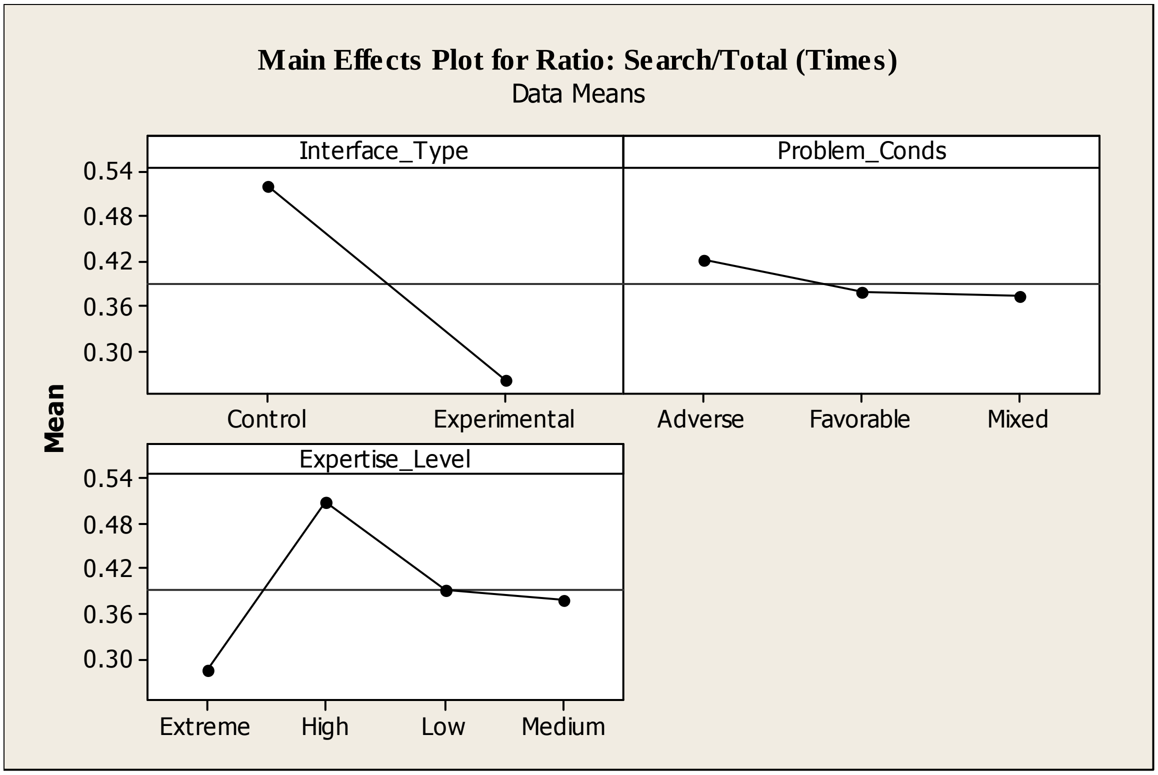


Figure 34: Main effects plot for "Ratio: Search/Total (Times)"

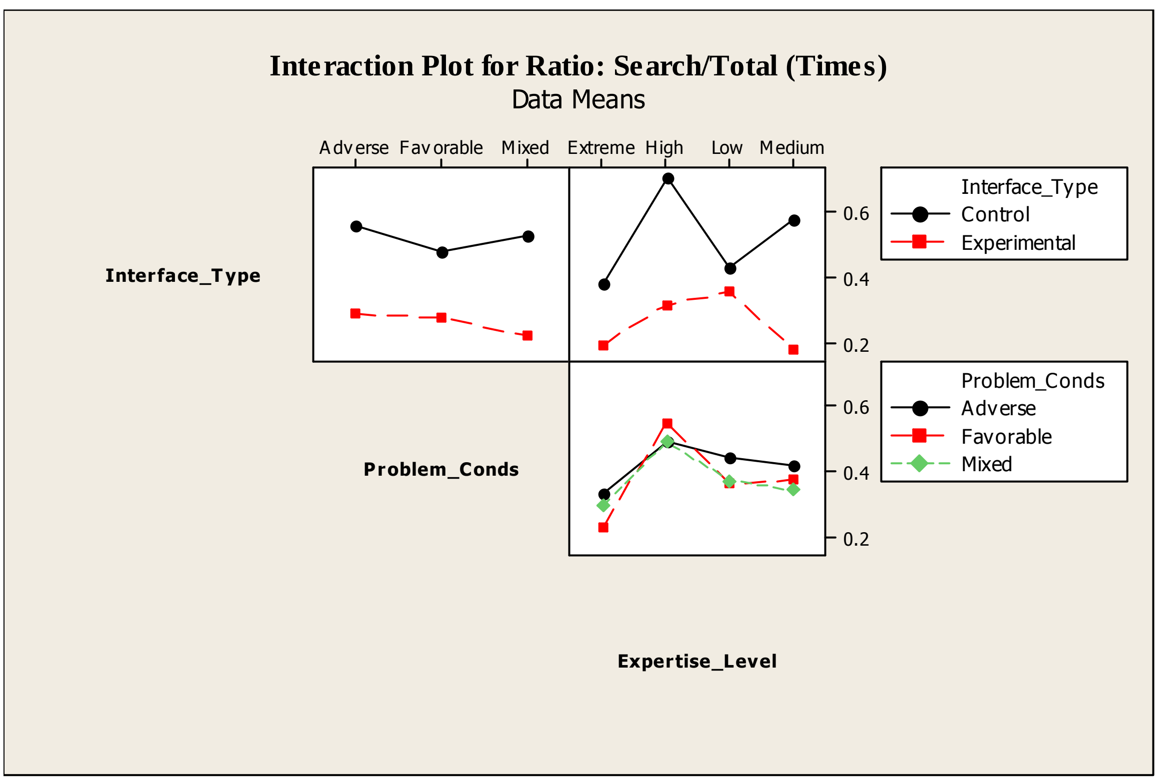


Figure 35: Interaction effects plot for "Ratio: Search/Total (Times)"

Ratio: Search/Total (Times)
Interface_Type, Problem_Conditions, Expertise_Level

Factor	Type	Levels	Values
Interface_Type	fixed	2	Control, Experimental
Problem_Conds	fixed	3	Adverse, Favorable, Mixed
Expertise_Level	fixed	4	Extreme, High, Low, Medium

Analysis of Variance for Ratio Search/Total, using AdjSS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface_Type	1	0.405864	0.405864	0.405864	**	
Problem_Conds	2	0.010872	0.010872	0.005436	**	
Expertise_Level	3	0.151961	0.151961	0.050654	**	
Interface_Type*Problem_Conds	2	0.010045	0.010045	0.005022	**	
Interface_Type*Expertise_Level	3	0.113444	0.113444	0.037815	**	
Problem_Conds*Expertise_Level	6	0.017391	0.017391	0.002898	**	
Interface_Type*Problem_Conds* Expertise_Level	6	0.025269	0.025269	0.004212	**	
Error	0	*	*	*		
Total	23	0.734846				

Figure 36: ANOVA table for "Ratio: Search/Total (Times)"

7.2.3.1 Recognition of main effects

From Figure 34 we observe the presence of main effects since the treatment means curves have all non-zero slopes. This result is consistent with the observations made on the "Documents Search Times" response; however, in this case, the marginal effects of 'Interface type' – although still the most significant – are not the only ones. We observe that the normalization process of this response unveiled the importance of the 'Expertise Level' factor. From Figure 36, we observe the magnitude of these factors as: SS = 0.405 (%SS = 55.2%) and SS = 0.152 (%SS = 20.7%) for 'Interface type' and 'Expertise Level', respectively. For a percentage comparison of the sum of squares concept refer to the discussion of results presented in the next section (Section 7.3).

7.2.3.2 Recognition of interaction effects

The comparison of interactions between factors for this response yields similar results to those found in the previous response, that is, there is a lack of parallelism in the treatment means curves (see Figure 35), which indicates the presence of interactions. An analysis of the sum of squares (see Figure 36) confirms the most significant interaction exists between ‘Interface type’ and ‘Expertise level’ (SS = 0.1134, %SS = 15.4%).

7.2.4 Analysis of factor effects for “Ratio: Used / Created (#Docs)”

The intention of this response is to measure the level of information overload for each treatment during the solution of the exercises. When a subject is solving an exercise he/she needs a certain number of documents; however, these documents are not known in advance. Subjects need to search for documents; during this search many documents are browsed or created, depending on the interface type. Many of those are not used for solving the problem. The smaller the ratio between documents used over documents searched/created, the smaller degree of information overload is experienced. The Treatment Means Table for this response is provided in Table 22. Using this table, we built Treatment Means Plots for main effects and Interaction effects; these are presented in Figure 37 and Figure 38, respectively. The Analysis of Variance (ANOVA) table for this response is presented in Figure 39. A subjective examination of the marginal and interaction effects using these plots and the ANOVA table completes this subsection.

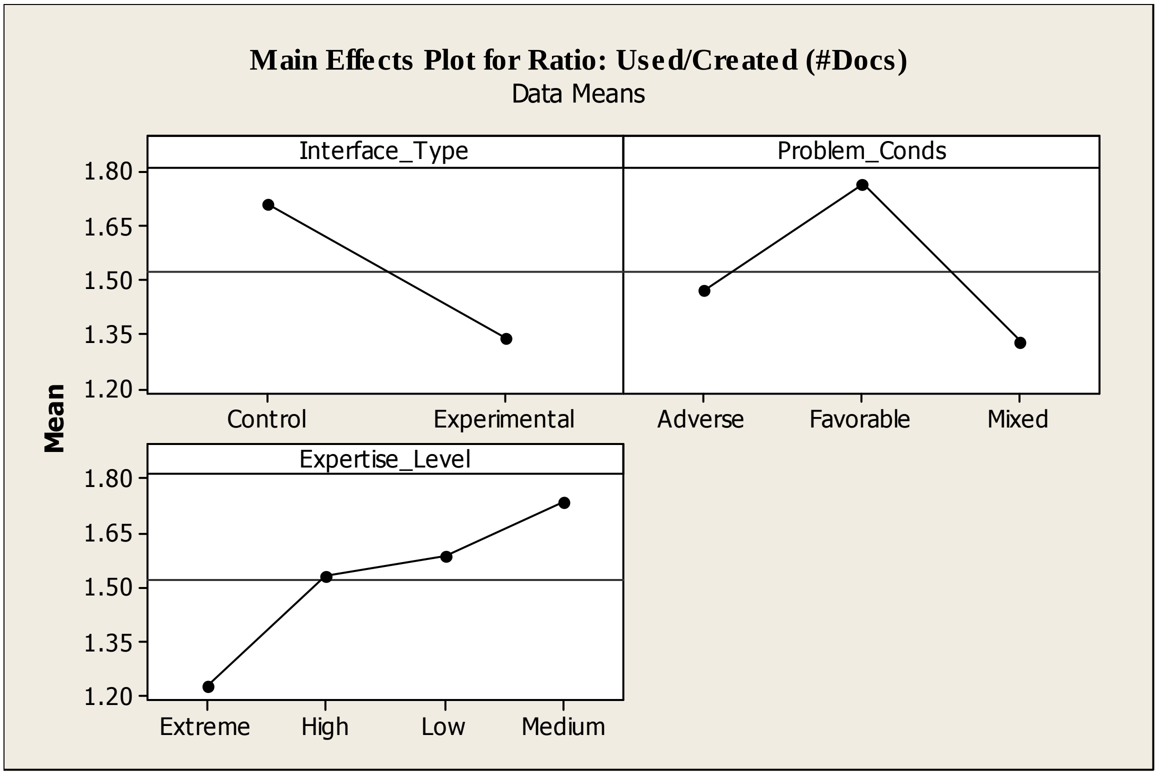


Figure 37: Main effects plot for "Ratio: Used/Created (#Docs)"

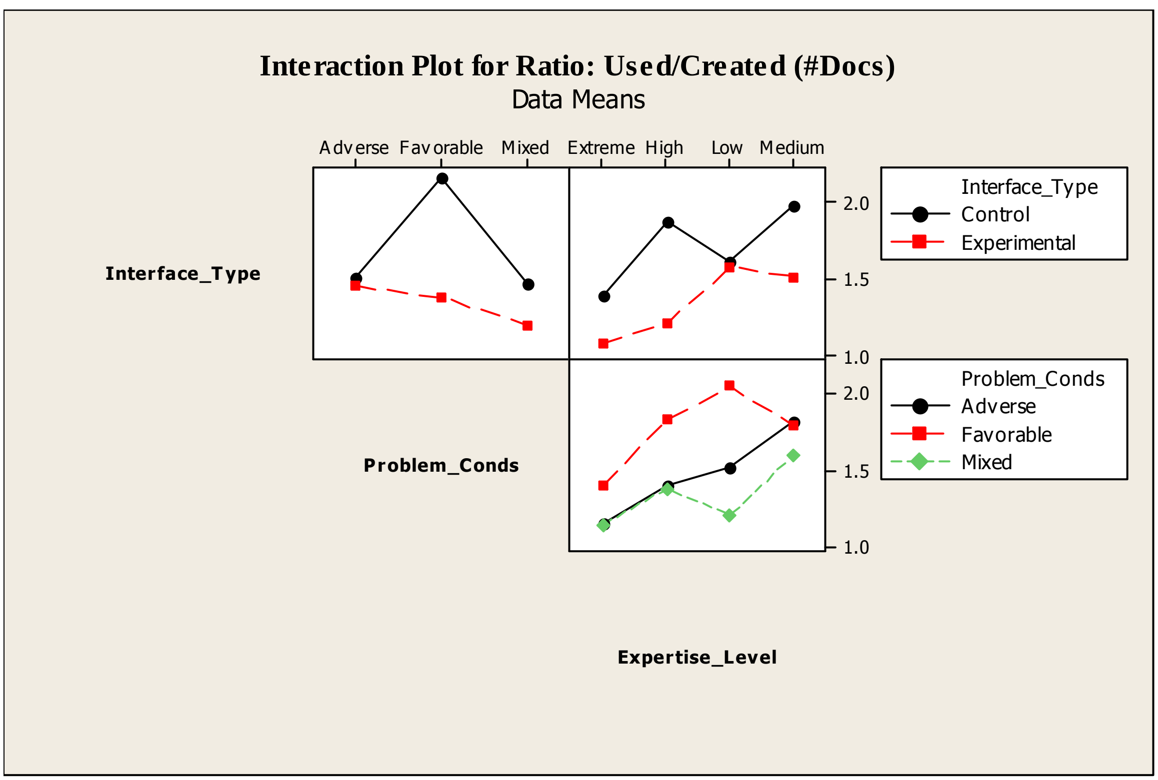


Figure 38: Interaction effects plot for "Ratio: Used/Created (#Docs)"

Ratio: Used/Created (#Docs)
Interface_Type, Problem_Conditions, Expertise_Level

Factor	Type	Levels	Values
Interface_Type	fixed	2	Control, Experimental
Problem_Conds	fixed	3	Adverse, Favorable, Mixed
Expertise_Level	fixed	4	Extreme, High, Low, Medium

Analysis of Variance for Ratio Used/Created, using SS for Tests

Source	DF	Seq SS	AdjSS	Adj MS	F	P
Interface_Type	1	0.82868	0.82868	0.82868	**	
Problem_Conds	2	0.78399	0.78399	0.39200	**	
Expertise_Level	3	0.81551	0.81551	0.27184	**	
Interface_Type*Problem_Conds	2	0.55972	0.55972	0.27986	**	
Interface_Type*Expertise_Level	3	0.32070	0.32070	0.10690	**	
Problem_Conds*Expertise_Level	6	0.33781	0.33781	0.05630	**	
Interface_Type*Problem_Conds* Expertise_Level	6	0.08150	0.08150	0.01358	**	
Error	0	*	*	*		
Total	23	3.72791				

Figure 39: ANOVA table for “Ratio: Used/Created (#Docs)”

7.2.4.1 Recognition of main effects

From a visual inspection of Figure 37 we observe that each of the treatment means curves has a non-zero slope, which indicates the presence of marginal effects of all the experimental factors. An inspection of the Sum of Squares in Figure 39 indicates the magnitude of these effects. As opposed to the previous responses, with this we notice that the three factors have similar effects in magnitude, i.e., ‘Interface Type’ has SS = 0.828 (%SS = 22.2%), ‘Problem Conditions’ SS = 0.783 (%SS = 21.0%), and ‘Expertise Level’ has SS = 0.815 (%SS = 21.9%). For a percentage comparison of the sum of squares concept refer to the discussion of results presented in the next section (Section 7.3).

7.2.4.2 Recognition of interaction effects

The comparison of interactions between factors for this response yields results similar to those found in the previous responses, that is, there is a lack of parallelism in the treatment means curves (see Figure 38), which indicates the presence on interactions. But yet, just like the type of marginal effects, the interaction effects for this response are more balanced than those of other responses. An analysis of the sum of squares (see Figure 39) confirms AC (‘Interface type’ and ‘Expertise level’) as the most significant interaction (SS = 0.559, %SS = 15.0%). AB (‘Interface type’ and ‘Problem Conditions’) and BC (‘Problem Conditions’ and ‘Expertise level’) interactions are slightly less important in magnitude: SS = 0.321 (%SS = 8.6%) and SS = 0.338 (%SS = 9.1%), for AB and BC interactions respectively.

7.2.5 Analysis of factor effects for “Total Costs ($)”

The “Total Cost ($)” response is a cost-based performance measure associated with the numeric solution obtained in each treatment. As explained in Chapter 6, the total cost comprises of three components; these reflect the costs for excess inventory, the acquisition of raw material and the loss of opportunity. A smaller total cost is desirable, therefore, smaller amounts will be considered as more beneficial. The Treatment Means Table for this response is provided in Table 23. Using this table, we built Treatment Means Plots for main effects and Interaction effects; these are presented in Figure 40 and Figure 41, respectively. The Analysis of Variance (ANOVA) table for this response is presented in Figure 42. A subjective examination of the marginal and interaction effects using these plots and the ANOVA table complete this subsection.

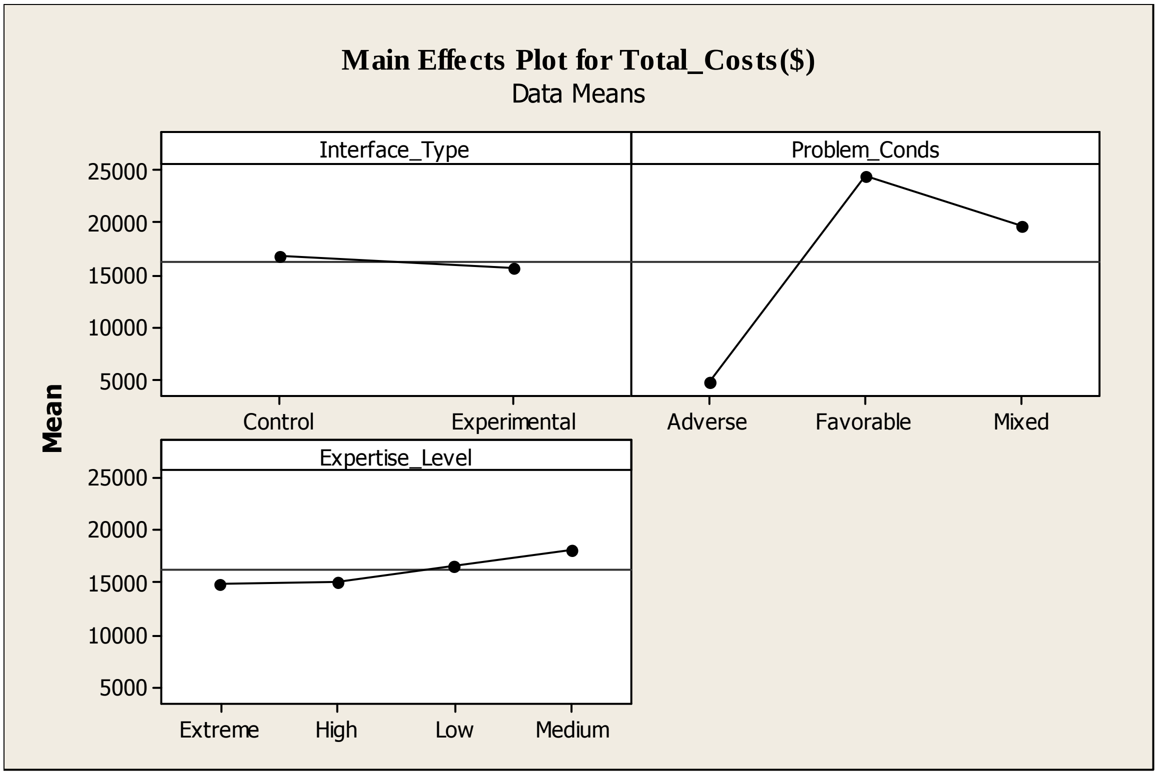


Figure 40: Main effects plot for "Total Costs ($)"

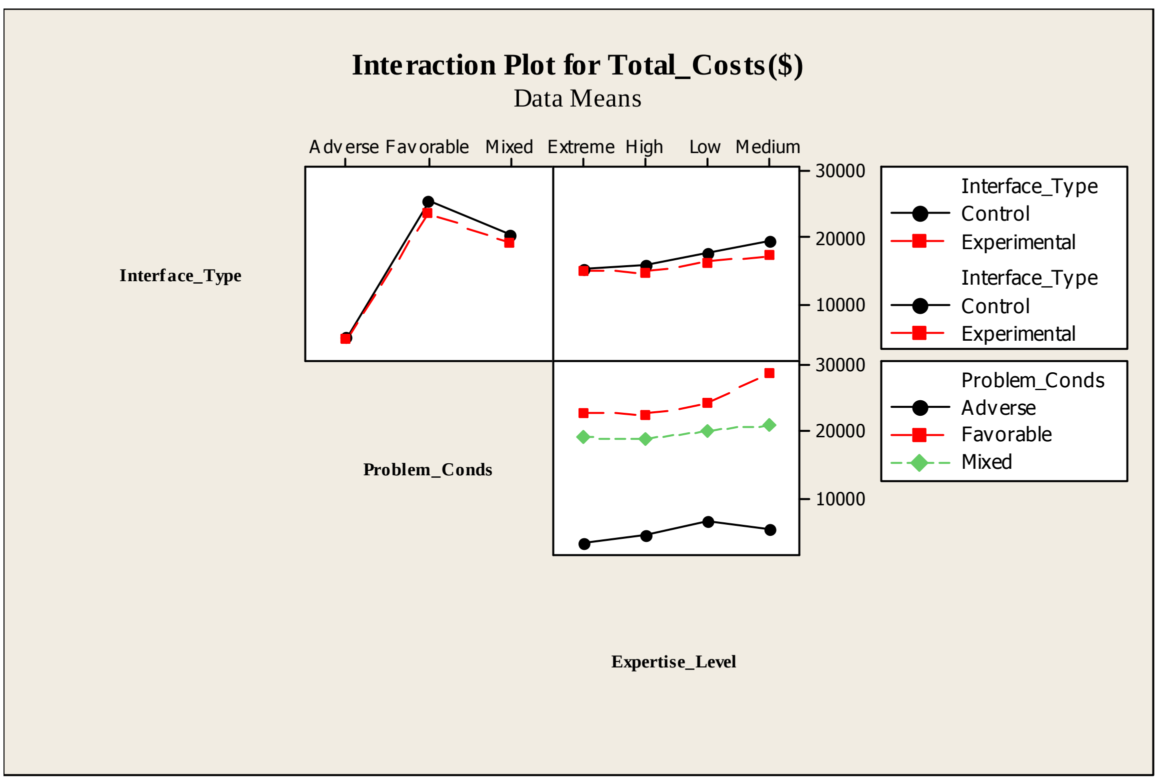


Figure 41: Interaction effects plot for "Total Costs ($)"

Total_Costs ($)
Interface_Type, Problem_Conds, Expertise_Level

Factor	Type	Levels	Values
Interface_Type	fixed	2	Control, Experimental
Problem_Conds	fixed	3	Adverse, Favorable, Mixed
Expertise_Level	fixed	4	Extreme, High, Low, Medium

Analysis of Variance for Total_Costs($), using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface_Type	1	8640000	8640000	8640000	**	
Problem_Conds	2	1698446858	1698446858	849223429	**	
Expertise_Level	3	44074953	44074954	14691651	**	
Interface_Type*Problem_Conds	2	2940393	2940393	1470196	**	
Interface_Type*Expertise_Level	3	2792769	2792769	930923	**	
Problem_Conds*Expertise_Level	6	23883025	23883025	3980504	**	
Interface_Type*Problem_Conds* Expertise_Level	6	5111491	5111491	851915	**	
Error	0	*	*	*		
Total	23	1785889490				

Figure 42: ANOVA table for “Total Costs ($)”

Table 23: Numerical results – Treatment Means Tables

Mean Total Costs according to Interface type, Problems' conditions, and Expertise level

Interface type (Factor A)	**Subjets' Expertise Level (Factor C) and Problem's Conditions (Factor B)**																			
	k=1; Extreme				k=2; High				k=3; Medium				k=4; Low				Average			
	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average	j=1; Favorable	j=2; Adverse	j=3; Mixed	Average
i =1; Control(C)	22,594 (μ_{111})	3,141 (μ_{121})	19,338 (μ_{131})	15,024 ($\mu_{1\cdot 1}$)	23,067 (μ_{112})	4,500 (μ_{122})	19,414 (μ_{132})	15,660 ($\mu_{1\cdot 2}$)	31,291 (μ_{113})	5,210 (μ_{123})	21,476 (μ_{133})	19,326 ($\mu_{1\cdot 3}$)	24,982 (μ_{114})	6,581 (μ_{124})	20,821 (μ_{134})	17,461 ($\mu_{1\cdot 4}$)	25,484 ($\mu_{11\cdot}$)	4,858 ($\mu_{12\cdot}$)	20,262 ($\mu_{13\cdot}$)	16,868 ($\mu_{1\cdot\cdot}$)
i =2; Experimental (E)	22,724 (μ_{211})	3,022 (μ_{221})	18,560 (μ_{231})	14,769 ($\mu_{2\cdot 1}$)	21,860 (μ_{212})	4,010 (μ_{222})	17,959 (μ_{232})	14,610 ($\mu_{2\cdot 2}$)	26,128 (μ_{213})	5,174 (μ_{223})	20,180 (μ_{233})	17,161 ($\mu_{2\cdot 3}$)	23,174 (μ_{214})	6,102 (μ_{224})	19,212 (μ_{234})	16,163 ($\mu_{2\cdot 4}$)	23,472 ($\mu_{21\cdot}$)	4,577 ($\mu_{22\cdot}$)	18,978 ($\mu_{23\cdot}$)	15,675 ($\mu_{2\cdot\cdot}$)
Average	22,659 ($\mu_{\cdot 11}$)	3,082 ($\mu_{\cdot 21}$)	18,949 ($\mu_{\cdot 31}$)	14,897 ($\mu_{\cdot\cdot 1}$)	22,464 ($\mu_{\cdot 12}$)	4,255 ($\mu_{\cdot 22}$)	18,687 ($\mu_{\cdot 32}$)	15,135 ($\mu_{\cdot\cdot 2}$)	28,710 ($\mu_{\cdot 13}$)	5,192 ($\mu_{\cdot 23}$)	20,828 ($\mu_{\cdot 33}$)	18,243 ($\mu_{\cdot\cdot 3}$)	24,078 ($\mu_{\cdot 14}$)	6,342 ($\mu_{\cdot 24}$)	20,017 ($\mu_{\cdot 34}$)	16,812 ($\mu_{\cdot\cdot 4}$)	24,478 ($\mu_{\cdot 1\cdot}$)	4,718 ($\mu_{\cdot 2\cdot}$)	19,620 ($\mu_{\cdot 3\cdot}$)	16,272 ($\mu_{\cdot\cdot\cdot}$)

Table 24: Analysis of Sum of Squares (Total and Percentage)

Source		**Total Execution Times**		**Documents Search (Times)**		**Ratio: Search / Total (Times)**		**Ratio: Used/Created (#Docs)**		**Total_Costs ($)**	
		Adj SS	**%SS**	**Adj SS**	**%SS**	**Adj SS**	**%SS**	**Adj SS**	**%SS**	**Adj SS**	**%SS**
A	Interface_Type	3457.24	32.5%	8415.51	70.3%	0.405864	55.2%	0.828680	22.2%	8640000	0.5%
B	Problem_Conds	1270.92	11.9%	663.37	5.5%	0.010872	1.5%	0.783990	21.0%	1698446858	95.1%
C	Expertise_Level	5044.60	47.4%	558.25	4.7%	0.151961	20.7%	0.815510	21.9%	44074953	2.5%
AB	Interface_Type*Problem_Conds	16.67	0.2%	112.10	0.9%	0.010045	1.4%	0.559720	15.0%	2940393	0.2%
AC	Interface_Type*Expertise_Level	413.20	3.9%	1291.85	10.8%	0.113444	15.4%	0.320700	8.6%	2792769	0.2%
BC	Problem_Conds*Expertise_Level	226.65	2.1%	382.66	3.2%	0.017391	2.4%	0.337810	9.1%	23883025	1.3%
ABC	Interface_Type*Problem_Conds*Expertise_Level	206.85	1.9%	539.43	4.5%	0.025269	3.4%	0.081500	2.2%	5111491	0.3%
	Error	*		*		*		*		*	
	Total	10636.13	100.0%	11963.17	100.0%	0.734846	100.0%	3.727910	100.0%	1785889489	100.0%

7.2.5.1 Recognition of main effects

From Figure 40, it can be seen that each of the treatment means curves has a non-zero slope, which indicates the presence of marginal effects of all the experimental factors. With a visual inspection it is noticeable that the treatment means curve for the ‘Problem Conditions’ factor presents the largest differential in heights. This fact is confirmed using the sum of squares concept, which provides a quantitative insight of the degree of importance of each marginal effect. From Figure 42 we can see that the ‘Problem Conditions’ constitutes the most significant factor (SS = 1698446858, %SS=95.1%). Other effects, although present, are less important in magnitude, i.e., ‘Interface Type’ has SS = 8640000 (%SS=0.5%) and ‘Expertise Level’ has SS = 44074953 (%SS=2.5%). These observations are consistent with the fact that treatment means curves for factors A and C are nearly horizontal. For a percentage comparison of the sum of squares concept refer to the discussion of results presented in the next section (Section 7.3).

7.2.5.2 Recognition of interaction effects

A visual inspection of whether the treatment means curves for the different factor levels in a treatment means plot are parallel is helpful to recognize the presence of interactions. Figure 41 presents the interactions plot for this response. We observe that interactions are of small magnitude. An analysis of the sum of squares Figure 42 confirms BC (‘Problem Conditions’ and ‘Expertise level’) as the most significant interaction (SS = 23883025, %SS = 1.3%). Other interactions AB (‘Interface type’ and ‘Problem Conditions’) and AC (‘Interface type’ and ‘Expertise level’) are slightly less important in magnitude, in both cases %SS = 0.2%.

7.3 Discussion of results

In this section we present a discussion of the results obtained during the empirical evaluation of the decision aid. The subjective observations made during the analysis of marginal and interaction effects of experimental factors are summarized in the discussion. The subjective assessments are supported with quantitative observations of the sum of squares (total and percentage) results presented in Table 24.

Before we initiate the discussion we need to recall the goals of the model implementation. The model provides a structure that integrates qualitative and quantitative data and that facilitates access to relevant data to aid the decision making process. An evaluation of the model implementation addresses the benefits achieved. In the particular case of the domain of application, we proposed transactional and numerical measures. In order to analyze the level at which these goals are achieved we study the speed with which data is accessed. The faster the data are accessed, the fewer are the amounts of unnecessary data that are accessed. Since the decision maker is subject to less data overload, he/she can perform better in his/her decision making. Three types of responses are being analyzed: (i) responses that measure speed; (ii) responses that support for data overload, and (iii) more accuracy on decisions (better solutions).

We intend to answer the first question, i.e., whether or not faster interactions are achieved and consequently whether or not data are accessed faster with the first three responses (Total Execution times – M1.1, Documents Search Times – M1.2 and Ratio: Search/total – M1.3). By analyzing the percentage of sum of squares (%SS) for these measures in Table 24 we observe that the factor A (‘Interface Type’) plays a significant role in the achievement of faster responses. For instance, if we observe marginal effects of factor A, for “Total Execution Times” it has a significant importance ranking in 2nd place (%SS = 32.5%), for “Document Search Times” it ranks in 1st place (70.3%) and for “Ratio: Search/Total” it ranks again first (55.2%). With respect to interaction effects, we notice that the most noticeable interaction exists with respect to factor C (‘Expertise Level’). For the three performance measures under study, these interactions are: 3.9%, 10.8%, and 15.4%. This is not surprising, since the marginal effects of factor C are noticeable for “Total Execution Times” (47.7%) and “Ratio: Search/Total” (20.7%). In conclusion, we observe that the ‘Interface Type’ in conjunction with the ‘Expertise Level’ is a significant factor in achieving faster access to required data. Next, we analyze the marginal and interaction effects of factors for the achievement of less data overload.

From Table 24 we observe that the three factors A (‘Interface Type’), B (‘Problem Conditions’), and C (‘Expertise Level’) have similar marginal effects on the achievement of low data overload (%SS = 21%); however, the interaction effects again reveal a significant AC interaction (%SS = 15%). Once again, we notice that ‘Interface Type’ has noticeable marginal and interaction effects for the achievement of less data overload. Next, we analyze the marginal and interaction effects of factors for the achievement of better decisions (more accuracy).

From Table 24 we observe that factor A (‘Interface Type’) has a minimal marginal effect on response (%SS = 0.5%); with respect to marginal effects AB and AC interactions are also minimal. The implication of this is that the factor of interest (‘Interface Type’) has very small effect on the achievement of better solutions. We attempt to provide an explanation to this fact next.

As we have seen, the achievement of faster access to data has a strong relationship with the type of interface as well as the expertise level and the combination of both factors. Similar observations can be made for the level of data overload observed. The achievement of these two goals leads us to conclude that the modeling methodology and the software implementation ("Interface Type" factor) are in fact providing the subjects with the necessary means to structure data (reduction in time to perform a solution). The interface Type factor also facilitates access to required data (marginal and interaction effects to reduce data overload). However, the increase in execution speed and reduction of data overload do not necessarily lead automatically to better decisions, in the case of the domain of application, to more accurate solutions. We could have anticipated these results give the important effect that 'Expertise Level' factor has on the experimental responses.

An immediate consequence drawn from the previous observations is the need to include in the experimental interface specific support for utilizing the required data and not just for accessing them. This addition would reflect in improved performance for decision making output. More about this is discussed next in the concluding chapter.

CHAPTER 8

CONTRIBUTIONS AND FUTURE EXTENSIONS

In this dissertation we have addressed the problem of modeling expertise in domains characterized by unquantifiable, often subjective, information, and using that model of expertise as the foundation for building computer-based decision support systems. The key feature of the expert model is to make explicit the essential characteristics of the knowledge experts use to process objective, quantitative information, for making decisions in environments rich in qualitative data. This model is then used as the basis for an “intelligent” interactive assistant that presents information appropriate for the context to operators who may not have developed the necessary expertise.

The core of the assistant is a heuristic algorithm that reflects what an expert decision maker would actually do. The algorithm incorporates a set of production rules, i.e., if-then-else rules, to define relevance conditions of quantitative data. These rules employ a *dominance principle*, i.e., a heuristic association of the relevance of quantitative data with the attributes of qualitative data, characterized as a set of ordered values. The heuristic algorithm is embedded in the assistant and is used to assist non-expert operators in locating information useful for making decisions.

The modeling methodology and the heuristic algorithm are applicable for modeling expertise in a class of decision problems characterized by large amounts of qualitative and quantitative data. The process of structuring the expert’s knowledge requires empirical evidence from actual decision problems; this evidence feeds the algorithm with heuristic associations between qualitative and quantitative data. The algorithm uses the dominance principle to decide what information to present for a particular set of conditions.

A summary of the methodology and details of the research contributions are presented next. In the final section of this chapter, future extensions of this work are proposed.

8.1 Summary of the modeling methodology

A methodology to structure decision problems characterized by large amounts of qualitative and quantitative data has been developed. The latter can often be processed and analyzed using mathematical or computational models. Qualitative data, on the other hand, need to be transformed to make them compatible with quantitative information. Typically, methodologies such as multi-attribute utility analysis are used to transform qualitative, subjective, data. However, the use of such methodologies assumes that qualitative data can easily be mapped and integrated with relevant quantitative data. This is rarely the case in practice, when large information systems are used to store and provide access to quantitative data.

In most contexts, quantitative data are easy to understand since they include all objective and directly measurable facts. However, it is often difficult, and not straightforward, to make sense of qualitative data, because such data refer to subjective assessments, interpretations, and judgments about external factors influencing a decision problem. Examples of these include: market uncertainty, unknown and uncertain causes for a disease, unknown and unquantifiable risks for adopting a certain policy, etc. In such environments, expert decision makers transform subjective, qualitative data into usable forms even when objective measures to represent these data are not readily available. A modeling methodology to characterize this expert knowledge has been developed.

The objective of the research described in this dissertation is to model decision problems by capturing and structuring the expertise needed to process qualitative and quantitative data. An expert decision maker *transforms* qualitative data in the form of ordinal quantities. This transformation facilitates the assessment of the qualitative characteristics of a given scenario. An expert also creates heuristic *associations* between qualitative and quantitative data. These associations are then used by the decision aid to identify needed quantitative data. The expert searches and accesses required quantitative data for making decisions. The model *structures* the information space using a directed graph.

A model of these processes *(transformation, association, and structure)* provides the foundation necessary to develop a heuristic algorithm that incorporates the relevance conditions of quantitative data in the form of production rules, i.e., if-then-else rules. These rules employ a *dominance principle,* i.e., a heuristic association of the relevance of quantitative data with the attributes of qualitative data, characterized as a set of ordered values. The heuristic algorithm is embedded in the assistant and is used to assist non-expert operators in locating information useful for making decisions.

8.2 Research contributions

We have developed a novel approach to heuristically associate the relevance of quantitative data with qualitative data for making decisions. Empirical evidence showed that attributes of qualitative data can be represented as an ordered scale of values and that the relevance of quantitative data can be associated with a specific range of the ordered scale following a *dominance principle*.

The methodology for modeling expertise as three independent and non-sequential processes is also new. When modeling the process of *transforming* qualitative data into a quantitative assessment, the model represents the influencing external factors and their attributes within an ordinal scale or an ordered representation. This technique is somewhat analogous to that of traditional methodologies such as multi-attribute utility analysis in which qualitative attributes are quantified during the problem representation.

The *association* of the values assessed by the expert for qualitative data with specific quantitative data forms a key component of the decision aid that projects a particular configuration of external influences into a specific set of needed quantitative data. The decision aid uses a directed graph to *structure* the information space necessary to search and access the required quantitative data.

Finally, the computational system implemented for visualizing relationships between qualitative and quantitative data offers a novel approach to information visualization.

8.3 Future extensions

In this dissertation, we described an empirical evaluation of the software implementation to test the effectiveness of the graph-based modeling approach for addressing production planning decisions. During the evaluation we encountered a limitation in the number of subjects available to perform the tasks under study. As mentioned before, the required expertise was such that only a few subjects were available. This fact limited the number of observations to only one per experimental treatment. Because of this, the analysis of results was limited to subjective observations of the differences between treatment means. In order to address this situation, future extensions to this work should increase the number of subjects. That would permit the testing of hypothesis about significant differences among marginal and interaction effects of experimental factors.

Another opportunity for future extensions is related to the accuracy of solutions obtained during each treatment. The modeling methodology is currently yielding positive results in terms of speeding up the retrieval of information; however, there is still margin for further improvements in numeric accuracy of solutions. Based on this, the *miniERP-GDSS* implementation is amenable to growing its functionalities as a decision support system. The inclusion of other data analysis features will certainly provide support for decision makers to make better use of relevant data.

Finally, new implementations of the modeling methodology in different domains are required to test the generality of the modeling approach. New implementations may broaden the *miniERP-GDSS*, or implement decision support systems in other domains, e.g., health care, stock investments, etc. Domains characterized by dynamic environments and IT-based systems to store quantitative information are likely to benefit from using the graph-based modeling methodology together with appropriately modified heuristic algorithms.

APPENDIX A.

The manufacturing operations at the CCV plant

In this appendix we provide a detailed description of the manufacturing processes at the Cooper Cameron Valves (CCV) manufacturing plant. The CCV plant is a metal-manufacturing facility that produces a variety of products used in the petroleum industry. It has annual sales of approximately 500 million dollars. The products discussed in this dissertation are ball valves used in the gas and oil industry.

Ball valve products

The CCV plant produces more than 25 different types of ball valves. Figure 43 shows a typical ball valve and its main components. According to the product specifications and manufacturing processes, total ball valve production is categorized in four production lines. Each one has particular needs of raw material and sequence of manufacturing processes operations. Table 25 depicts the fours production families.

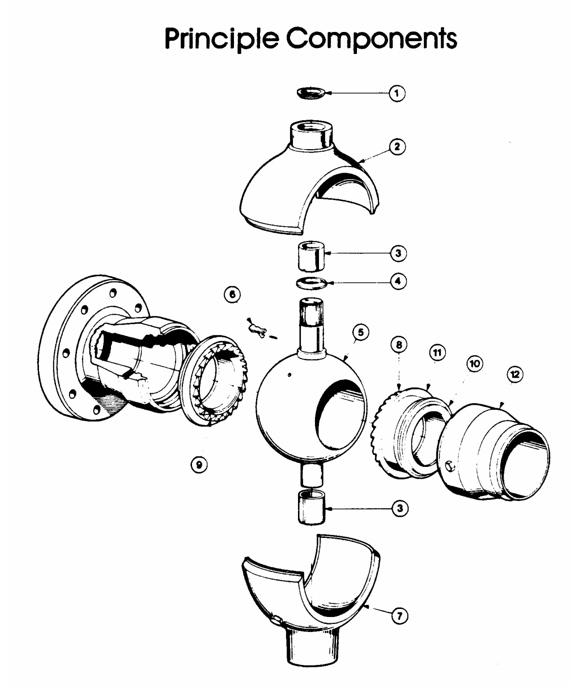


Figure 43: Ball Valve and its principle components

Product flow

As indicated above, each product group has its own routing sequence with individual instructions at each work center to assemble a final product. For example, when a production order containing a product of group 3 (e.g., a ball valve of 24" of diameter) is released for production on the factory floor, the enterprise resource planning (ERP) information system installed in the CCV plant generates a series of messages to advertise all the entities related to the manufacture of such product (raw materials, machines, material handling systems, etc.) the existence of the new item to be manufactured.

Table 25: Categories of the products manufactured at the CCV Plant

Group	Ball Valve Description	P/n
A	Ball valve 8" ANSI Class 150# RFxRF Trim 212 FP	08-01-01-F
	Ball valve 8" ANSI Class 300# RFxRF Trim 212 FP	08-02-01-F
	Ball valve 8" ANSI Class 600# RFxRF Trim 212 FP	08-03-01-F
	Ball valve 10" ANSI Class 150# RFxRF Trim 212 FP	10-01-01-F
	Ball valve 10" ANSI Class 300# RFxRF Trim 212 FP	10-02-01-F
	Ball valve 10" ANSI Class 600# RFxRF Trim 212 FP	10-03-01-F
B	Ball valve 12" ANSI Class 150# RFxRF Trim 212 FP	12-01-01-F
	Ball valve 12" ANSI Class 300# RFxRF Trim 212 FP	12-02-01-F
	Ball valve 12" ANSI Class 600# RFxRF Trim 212 FP	12-03-01-F
	Ball valve 16" ANSI Class 150# RFxRF Trim 212 FP	16-01-01-F
	Ball valve 16" ANSI Class 300# RFxRF Trim 212 FP	16-02-01-F
	Ball valve 16" ANSI Class 600# RFxRF Trim 212 FP	16-03-01-F
C	Ball valve 20" ANSI Class 150# RFxRF Trim 212 FP	20-01-01-F
	Ball valve 20" ANSI Class 300# RFxRF Trim 212 FP	20-02-01-F
	Ball valve 20" ANSI Class 600# RFxRF Trim 212 FP	20-03-01-F
	Ball valve 24" ANSI Class 150# RFxRF Trim 212 FP	24-01-01-F
	Ball valve 24" ANSI Class 300# RFxRF Trim 212 FP	24-02-01-F
	Ball valve 24" ANSI Class 600# RFxRF Trim 212 FP	24-03-01-F
D	Ball valve 30" ANSI Class 150# RFxRF Trim 212 FP	30-01-01-F
	Ball valve 30" ANSI Class 300# RFxRF Trim 212 FP	30-02-01-F
	Ball valve 30" ANSI Class 600# RFxRF Trim 212 FP	30-03-01-F
	Ball valve 36" ANSI Class 150# RFxRF Trim 212 FP	36-01-01-F
	Ball valve 36" ANSI Class 300# RFxRF Trim 212 FP	36-02-01-F
	Ball valve 36" ANSI Class 600# RFxRF Trim 212 FP	36-03-01-F

Generic process routing sequence

In this section we describe a generic routing sequence to manufacture a ball valve in the CCV plant. The component parts names and process steps mentioned refer to Figure 43 and Figure 44 respectively.

Prerequisites:

The manufacture of one product is subject to two conditions: (i) a production order has been released and scheduled; and (ii) the required components (raw materials or sub-assemblies) to produce the order are in stock.

Step 0

The inventory warehouse releases the following parts to three different workstations in the manufacturing shop floor: (i) lower and upper shells, trunnion, and end connections (flanges), (ii) seat rings kits, and (iii) balls.

Step 1A: Machining center

In this workstation the following parts are received: ‘Lower and upper shells’, ‘trunnion’, and ‘end connections’ (flanges). These parts are machined until design specifications are achieved. The machining center comprises two machines, a loading robot, an unloading robot, and two buffers to keep incoming and outgoing parts. When these parts have been processed, they are sent to different workstations. For example, when ‘trunnion’ has been machined it is sent to Ball-Trunnion Assembly Center (step 3A). When the upper and lower shells have been machined, they are sent to Shells-Trunnion Assembly Center (step 4A). Finally, the end connections are sent to the Flanges-Seats Assembly Center (step 6A).

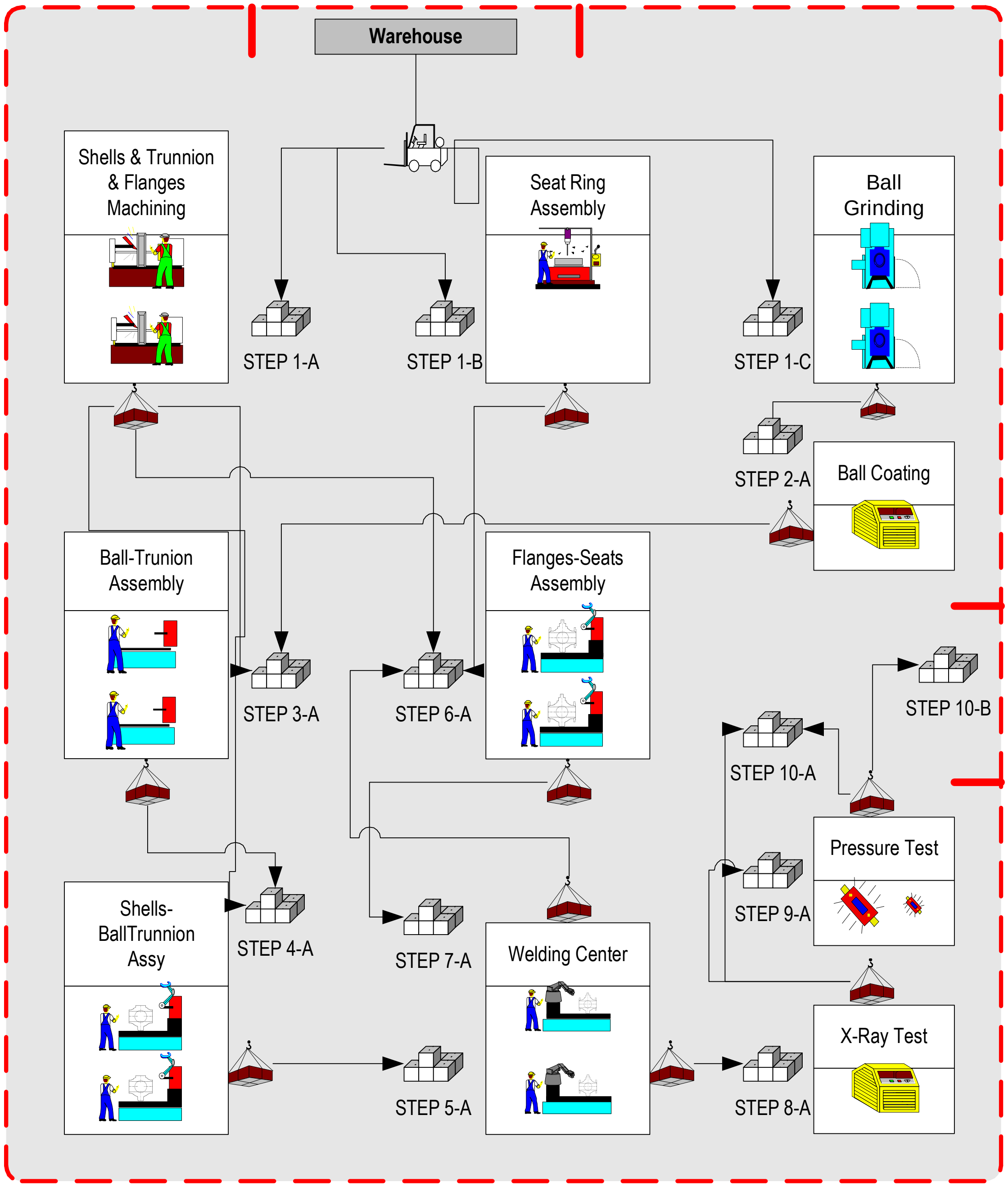


Figure 44: Factory floor and process routing sequence

Step 1B: Seat ring Assembly Center

The seat rings kits are assembled in this workstation. The kits are received directly from the inventory stock (step 0). This assembly center consists of: one machine, one mobile robot to load and unload parts in the machine, and two buffers to keep incoming and outgoing parts. When these kits have been assembled they are sent to the Flanges-Seats Assembly Center (step 6A).

Step 1C: Grinding Center

The balls are the closing component of the valves. They are sent to this work station. At this point, balls' preceding station is the inventory stock (step 0). The purpose of the grinding center is to provide a mirror surface to the ball. The grinding center comprises two grinding machines, a robot to load and unload parts, and two buffers to keep incoming and outgoing parts. After the grinding process, the ball is sent to the ball coating center (step 2), where it receives a special treatment.

Step 2A: Ball Coating Center

This workstation receives components coming from step 1C (Balls). The objective of this station is to apply an anticorrosion coating on the ball's surface. The workstation is comprised of two machines, a robot to load and unload parts, and two buffers to keep incoming and outgoing parts. Once the special coating has been received the balls are sent to the next workstation: step 3A.

Step 3A: Ball Trunnion Assembly Center

This workstation receives components from step 2A (Balls) and from step 1C (Trunnion). The objective of this station is to assemble these two parts into a single unit. The workstation has two machines, a robot to load and unload parts, and two buffers to keep incoming and outgoing parts. When the trunnion and the ball have been assembled, this unit is sent to the next station: step 4A.

Step 4A: Shells - Ball - Trunnion Assembly Center

This workstation receives components coming from step 3A (Ball – Trunnion) and from step 1A (upper and lower shells). Once the ball – trunnion unit has been released from the step 3A, it arrives at this workstation, where it waits until its correspondent upper and lower shells (that have been machined in step 1A) are released and sent to this workstation. When the three parts arrive they are assembled together. This subassembly is then sent to the next workstation (step 5A).

Step 5A: Welding Center

This workstation receives the subassembly coming from step 4A (Ball – Trunnion – Shells). The objective of this workstation is to join externally the shells that encapsulate the Ball – Trunnion using a welding procedure. The workstation is comprised of two welding robots, two manually operated robots to load and unload subassemblies, a robot to load and unload parts, and two buffers to keep incoming and outgoing parts. When the subassembly has been welded it is sent to the next workstation (step 6A).

Step 6A: Flanges – Seats – Subassembly Assembly Center

When the subassembly has been welded in step 5A it is then sent to this workstation where it will be added to two more components that arrive from other two workstations: End connections (Step 1A) and Seat Ring (Step 1B). These three components are assembled together and then sent back to the Welding Center (step 7A) where a final welding process to join these three components take place.

Step 7A: Welding Center

The welding center receives again the subassembly coming from step 6A but now added to all its internal components (Seats) and the end connections (flanges). The objective of this step 7A in the welding center is to join externally the subassembly and the end connections using a welding procedure. When the assembly has been welded it is sent to the next workstation (step 8A).

Step 8A: X-Ray Testing Center

The assembly's welded joints are tested in this workstation using an X-Ray Test procedure. If the assembly passes the test it is sent to next workstation (step 9A), if not then the assembly can be sent either to workstation in step 6A or to the welding center (step 7A).

The X-Ray Testing Center is composed of two buffers areas to keep incoming and outgoing parts, a closed room where parts receive X-Ray test, and two loaders to bring parts in and out the test room.

Step 9A: Pressure Testing Center

When the assembly has passed the X-ray test it comes to this workstation where it receives the final test. This test consists of pressurizing the valve to check the hermetic sealing. When the valve passes this second test it is finished and ready to ship.

APPENDIX B.

Expert knowledge inferred from ethnographic study

In this appendix we include the complete expert knowledge table inferred during the ethnographic study (presented in Chapter 3). The expert knowledge describes the relevance values for each data repository with respect to the attributes of each environmental factor.

Table 26: Expert knowledge inferred from ethnographic study

	Material Supply						Competition				Customer & Market						Management				Events	
Document Id	mtr01		mtr02		mtr03		cmp01		cmp02		mkt01		mkt02		mkt03		dmk01		dmk02		rdm01	
	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V
DF											**3**	**3**	**4**	**3**	**3**	**3**	**4**	**3**	**3**	**3**		
DF-001											3	0	4	1	3	3	4	3	3	3		
DF-002											3	1	4	1	3	2	4	2	3	3		
DF-003											3	1	4	1	3	1	4	2	3	0		
DF-004											3	3	4	3	3	3	4	2	3	3		
DF-005											3	3	4	2	3	3	4	1	3	2		
DF-006											3	2	4	0	3	3	4	2	3	1		
DF-007											3	2	4	1	3	3	4	1	3	3		
DF-008											3	1	4	2	3	1	4	3	3	2		
DF-009											3	1	4	1	3	2	4	1	3	1		
DF-010											3	2	4	2	3	1	4	0	3	1		
DF-011											3	2	4	1	3	1	4	1	3	2		
DF-012											3	3	4	3	3	3	4	2	3	3		
DF-013											3	0	4	3	3	3	4	2	3	2		
DF-014											3	2	4	1	3	3	4	2	3	3		
DF-015											3	2	4	1	3	3	4	1	3	1		
DF-016											3	3	4	1	3	3	4	3	3	1		
DF-017											3	2	4	3	3	0	4	2	3	3		
DF-018											3	1	4	2	3	1	4	2	3	2		
DF-019											3	2	4	3	3	1	4	3	3	2		
DF-020											3	3	4	0	3	1	4	3	3	1		
DF-021											3	2	4	2	3	1	4	2	3	3		
DF-022											3	2	4	3	3	2	4	3	3	1		
DF-023											3	3	4	2	3	3	4	2	3	3		
DF-024											3	3	4	2	3	2	4	1	3	2		
DF-025											3	1	4	3	3	2	4	2	3	1		
DF-026											3	2	4	2	3	3	4	1	3	2		
DF-027											3	2	4	3	3	3	4	3	3	1		
DF-028											3	3	4	3	3	2	4	2	3	2		
DF-029											3	2	4	3	3	1	4	0	3	1		
DF-030											3	2	4	3	3	1	4	3	3	1		

Table 26 (continued)

	Material Supply						Competition				Customer & Market						Management				Events	
Document	mtr01		mtr02		mtr03		cmp01		cmp02		mkt01		mkt02		mkt03		dmk01		dmk02		rdm01	
Id	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V
DF-031											3	1	4	2	3	3	4	1	3	1		
DF-032											3	2	4	2	3	1	4	2	3	1		
DF-033											3	3	4	2	3	1	4	2	3	1		
DF-034											3	1	4	2	3	1	4	1	3	3		
DF-035											3	2	4	3	3	1	4	2	3	2		
DF-036											3	3	4	2	3	2	4	1	3	1		
DF-037											3	2	4	3	3	2	4	3	3	2		
DF-038											3	2	4	1	3	1	4	1	3	3		
DF-039											3	1	4	1	3	2	4	1	3	3		
DF-040											3	1	4	1	3	3	4	2	3	3		
DF-041											3	1	4	3	3	3	4	3	3	3		
DF-042											3	1	4	1	3	3	4	2	3	2		
DF-043											3	1	4	2	3	2	4	3	3	3		
DF-044											3	3	4	1	3	1	4	2	3	1		
DF-045											3	3	4	3	3	2	4	3	3	3		
DF-046											3	3	4	0	3	3	4	3	3	1		
DF-047											3	2	4	3	3	1	4	1	3	1		
DF-048											3	3	4	2	3	3	4	2	3	3		
DF-049											3	1	4	1	3	1	4	3	3	1		
DF-050											3	1	4	1	3	2	4	1	3	2		
DF-051											3	1	4	3	3	1	4	3	3	3		
DF-052											3	3	4	1	3	2	4	3	3	1		
DF-053											3	1	4	1	3	3	4	1	3	3		
DF-054											3	3	4	2	3	3	4	1	3	3		
DF-055											3	2	4	3	3	3	4	3	3	3		
DF-056											3	1	4	2	3	0	4	2	3	3		
DF-057											3	2	4	3	3	2	4	2	3	2		
DF-058											3	3	4	1	3	2	4	3	3	1		
DF-059											3	3	4	2	3	3	4	1	3	2		
PH	**2**	**3**	**1**	**3**	**2**	**3**			**1**	**3**	**2**	**3**	**1**	**3**	**3**	**3**	**2**	**3**				
PH-001	2	2	1	1	2	1			1	3	2	3	1	3	3	1	2	1				
PH-002	2	1	1	2	2	2			1	3	2	2	1	2	3	2	2	2				
PH-003	2	2	1	3	2	1			1	2	2	3	1	1	3	3	2	2				
PH-004	2	3	1	3	2	3			1	3	2	3	1	1	3	3	2	2				
PH-005	2	1	1	1	2	3			1	1	2	1	1	1	3	2	2	3				
PH-006	2	2	1	1	2	1			1	2	2	3	1	1	3	3	2	1				
PH-007	2	3	1	1	2	2			1	2	2	2	1	1	3	2	2	3				
PH-008	2	1	1	2	2	3			1	3	2	1	1	1	3	1	2	1				
PH-009	2	1	1	2	2	1			1	1	2	3	1	3	3	1	2	3				
PH-010	2	2	1	1	2	3			1	2	2	0	1	2	3	3	2	1				
PH-011	2	2	1	0	2	2			1	1	2	3	1	2	3	1	2	2				
PH-012	2	3	1	2	2	3			1	3	2	3	1	3	3	1	2	1				
PH-013	2	2	1	2	2	2			1	1	2	3	1	1	3	3	2	3				
PH-014	2	2	1	3	2	1			1	1	2	2	1	1	3	2	2	2				
PH-015	2	3	1	3	2	3			1	1	2	2	1	2	3	3	2	1				
PH-016	2	1	1	1	2	3			1	2	2	2	1	0	3	1	2	3				
PH-017	2	1	1	2	2	2			1	3	2	2	1	1	3	2	2	3				
PH-018	2	3	1	1	2	1			1	3	2	3	1	3	3	1	2	3				
PH-019	2	3	1	3	2	2			1	2	2	1	1	3	3	3	2	2				
PH-020	2	2	1	3	2	0			1	1	2	3	1	1	3	1	2	2				

Table 26 (continued)

Document Id	Material Supply						Competition				Customer & Market						Management				Events	
	mtr01		mtr02		mtr03		cmp01		cmp02		mkt01		mkt02		mkt03		dmk01		dmk02		rdm01	
	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V
PH-021	2	1	1	1	2	1			1	3	2	2	1	1	3	2	2	1				
PH-022	2	0	1	3	2	1			1	1	2	3	1	2	3	1	2	3				
PH-023	2	1	1	1	2	1			1	3	2	2	1	1	3	1	2	3				
PH-024	2	2	1	2	2	2			1	3	2	2	1	2	3	3	2	3				
PH-025	2	3	1	2	2	1			1	3	2	1	1	2	3	1	2	1				
PH-026	2	1	1	1	2	1			1	3	2	1	1	2	3	2	2	1				
PH-027	2	1	1	3	2	2			1	3	2	2	1	1	3	1	2	2				
PH-028	2	3	1	2	2	1			1	3	2	3	1	3	3	2	2	2				
PH-029	2	2	1	2	2	1			1	2	2	3	1	1	3	3	2	2				
PH-030	2	1	1	2	2	2			1	3	2	3	1	1	3	3	2	1				
PH-031	2	2	1	2	2	2			1	1	2	3	1	3	3	1	2	1				
PH-032	2	1	1	2	2	3			1	2	2	1	1	2	3	1	2	1				
PH-033	2	2	1	1	2	3			1	1	2	1	1	3	3	2	2	2				
PH-034	2	3	1	1	2	3			1	2	2	2	1	2	3	2	2	1				
PH-035	2	0	1	1	2	1			1	1	2	1	1	2	3	2	2	1				
PH-036	2	2	1	1	2	2			1	2	2	3	1	2	3	3	2	3				
PH-037	2	3	1	1	2	2			1	3	2	2	1	3	3	1	2	1				
PH-038	2	3	1	1	2	1			1	1	2	3	1	1	3	2	2	3				
PH-039	2	2	1	1	2	3			1	2	2	3	1	3	3	3	2	2				
PH-040	2	2	1	3	2	1			1	2	2	1	1	3	3	1	2	1				
PH-041	2	2	1	1	2	3			1	2	2	0	1	3	3	1	2	2				
PH-042	2	2	1	3	2	3			1	2	2	3	1	3	3	3	2	2				
PH-043	2	2	1	3	2	2			1	1	2	2	1	1	3	1	2	2				
PH-044	2	2	1	2	2	3			1	2	2	3	1	1	3	2	2	3				
PH-045	2	1	1	3	2	3			1	3	2	2	1	1	3	1	2	3				
PH-046	2	2	1	1	2	2			1	1	2	2	1	1	3	3	2	2				
PH-047	2	2	1	3	2	2			1	1	2	1	1	3	3	2	2	2				
PH-048	2	2	1	3	2	2			1	1	2	3	1	0	3	1	2	2				
PH-049	2	2	1	3	2	3			1	1	2	3	1	2	3	3	2	1				
PH-050	2	1	1	3	2	3			1	3	2	3	1	1	3	1	2	2				
PB	**2**	**3**	**1**	**3**	**2**	**3**			**1**	**3**	**2**	**3**	**1**	**3**	**3**	**3**	**2**	**3**				
PB-001	2	3	1	3	2	1			1	3	2	1	1	2	3	2	2	3				
PB-002	2	3	1	3	2	1			1	3	2	1	1	2	3	2	2	3				
PB-003	2	1	1	3	2	2			1	3	2	3	1	1	3	1	2	1				
PB-004	2	1	1	2	2	3			1	2	2	0	1	2	3	1	2	3				
PB-005	2	1	1	2	2	1			1	2	2	1	1	3	3	2	2	1				
PB-006	2	1	1	3	2	3			1	3	2	1	1	2	3	2	2	1				
PB-007	2	3	1	1	2	2			1	1	2	3	1	1	3	3	2	1				
PB-008	2	1	1	1	2	1			1	2	2	3	1	2	3	2	2	1				
PB-009	2	3	1	2	2	1			1	2	2	1	1	2	3	3	2	2				
PB-010	2	1	1	3	2	1			1	3	2	2	1	2	3	2	2	2				
PB-011	2	3	1	1	2	2			1	2	2	3	1	1	3	1	2	2				
PB-012	2	3	1	3	2	2			1	3	2	3	1	2	3	1	2	3				
PB-013	2	2	1	3	2	2			1	1	2	3	1	2	3	2	2	2				
PB-014	2	3	1	1	2	2			1	3	2	1	1	2	3	2	2	2				
PB-015	2	3	1	1	2	3			1	1	2	2	1	2	3	2	2	3				
PB-016	2	2	1	2	2	3			1	0	2	2	1	1	3	3	2	2				
PB-017	2	2	1	1	2	3			1	3	2	2	1	3	3	3	2	1				
PB-018	2	3	1	1	2	3			1	1	2	1	1	3	3	2	2	1				
PB-019	2	0	1	1	2	1			1	0	2	3	1	3	3	3	2	1				

Table 26 (continued)

	Material Supply						Competition				Customer & Market						Management				Events	
Document Id	mtr01		mtr02		mtr03		cmp01		cmp02		mkt01		mkt02		mkt03		dmk01		dmk02		rdm01	
	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V
PB-020	2	1	1	1	2	2			1	3	2	3	1	1	3	1	2	1				
PB-021	2	3	1	3	2	3			1	1	2	1	1	2	3	1	2	2				
PB-022	2	1	1	1	2	1			1	3	2	1	1	3	3	1	2	2				
PB-023	2	3	1	3	2	2			1	2	2	3	1	2	3	3	2	3				
PB-024	2	1	1	2	2	3			1	3	2	1	1	3	3	2	2	3				
PB-025	2	2	1	2	2	0			1	1	2	1	1	3	3	3	2	3				
SC			**1**	**3**			**4**	**3**	**4**	**3**	**4**	**3**	**4**	**3**	**2**	**3**	**3**	**3**	**2**	**3**	**2**	**3**
SC-001			1	2			4	1	4	2	4	0	4	3	2	2	3	1	2	2	2	1
SC-002			1	1			4	3	4	2	4	3	4	1	2	2	3	2	2	3	2	1
SC-003			1	1			4	2	4	3	4	3	4	3	2	2	3	3	2	0	2	2
SC-004			1	3			4	1	4	1	4	2	4	3	2	1	3	2	2	1	2	2
SC-005			1	3			4	1	4	2	4	1	4	1	2	0	3	3	2	3	2	2
SC-006			1	2			4	3	4	1	4	3	4	2	2	2	3	1	2	2	2	1
SC-007			1	3			4	2	4	2	4	3	4	1	2	1	3	3	2	1	2	1
SC-008			1	2			4	2	4	3	4	1	4	2	2	3	3	1	2	1	2	2
SC-009			1	3			4	1	4	2	4	1	4	1	2	1	3	2	2	2	2	2
SC-010			1	3			4	2	4	1	4	3	4	1	2	1	3	3	2	2	2	3
SC-011			1	2			4	3	4	3	4	1	4	1	2	1	3	3	2	3	2	1
SC-012			1	2			4	2	4	3	4	2	4	2	2	0	3	3	2	2	2	3
SC-013			1	1			4	3	4	2	4	3	4	2	2	2	3	2	2	2	2	2
SC-014			1	3			4	1	4	1	4	0	4	1	2	2	3	2	2	1	2	3
SC-015			1	2			4	3	4	1	4	1	4	2	2	2	3	1	2	2	2	1
SC-016			1	3			4	1	4	1	4	1	4	3	2	3	3	3	2	3	2	3
SC-017			1	2			4	2	4	1	4	1	4	3	2	3	3	2	2	3	2	3
SC-018			1	2			4	2	4	1	4	2	4	3	2	1	3	1	2	2	2	1
SC-019			1	2			4	2	4	3	4	1	4	3	2	1	3	2	2	3	2	3
SC-020			1	3			4	3	4	3	4	3	4	2	2	1	3	2	2	1	2	1
SC-021			1	1			4	0	4	2	4	1	4	2	2	3	3	3	2	2	2	1
SC-022			1	2			4	1	4	1	4	2	4	1	2	3	3	1	2	3	2	2
SC-023			1	2			4	3	4	2	4	1	4	1	2	2	3	2	2	2	2	1
SC-024			1	3			4	2	4	2	4	0	4	3	2	1	3	2	2	1	2	2
SC-025			1	1			4	3	4	3	4	2	4	1	2	1	3	3	2	2	2	1
SC-026			1	2			4	1	4	3	4	1	4	1	2	3	3	1	2	2	2	1
SC-027			1	2			4	2	4	1	4	2	4	3	2	3	3	2	2	1	2	3
SC-028			1	3			4	3	4	1	4	3	4	3	2	3	3	2	2	3	2	1
SC-029			1	1			4	2	4	3	4	1	4	3	2	1	3	1	2	3	2	3
SC-030			1	2			4	2	4	2	4	2	4	2	2	1	3	1	2	3	2	2
VD	**2**	**3**	**1**	**3**	**2**	**3**			**1**	**3**	**2**	**3**	**1**	**3**	**3**	**3**	**2**	**3**				
VD-001	2	2	1	3	2	1			1	3	2	1	1	1	3	3	2	1				
VD-002	2	1	1	2	2	3			1	1	2	1	1	2	3	1	2	3				
CO	**2**	**3**	**1**	**3**	**2**	**3**			**1**	**3**	**2**	**3**	**1**	**3**	**3**	**3**	**1**	**3**			**1**	**3**
CO-001	2	1	1	3	2	2			1	1	2	3	1	3	3	2	1	2			1	1
CO-002	2	1	1	1	2	3			1	1	2	2	1	2	3	1	1	1			1	1
CO-003	2	1	1	3	2	3			1	3	2	1	1	2	3	2	1	3			1	2
CO-004	2	3	1	3	2	0			1	3	2	2	1	1	3	0	1	3			1	2
CO-005	2	3	1	3	2	3			1	3	2	1	1	1	3	1	1	3			1	1
PC	**2**	**3**									**2**	**3**							**4**	**3**	**4**	**3**
PC-001	2	1									2	2							4	1	4	2
PC-002	2	2									2	1							4	3	4	1
PC-003	2	2									2	1							4	1	4	2

Table 26 (continued)

	Material Supply						Competition				Customer & Market						Management				Events	
Document Id	mtr01		mtr02		mtr03		cmp01		cmp02		mkt01		mkt02		mkt03		dmk01		dmk02		rdm01	
	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V	I	V
PC-004	2	3									2	2							4	2	4	2
PC-005	2	3									2	1							4	1	4	3
PC-006	2	1									2	2							4	2	4	2
PC-007	2	1									2	3							4	2	4	2
PC-008	2	3									2	1							4	1	4	1
PC-009	2	1									2	2							4	0	4	1
PC-010	2	2									2	2							4	3	4	3
PC-011	2	2									2	3							4	2	4	2
PC-012	2	1									2	1							4	3	4	2
PC-013	2	1									2	3							4	2	4	2

APPENDIX C.

Profile of subjects participating in the empirical evaluation

In this appendix we describe the profiles of the subjects that participated in the empirical evaluation. The profile is described in terms of six fields: (i) experience, (ii) expertise level, (iii) role, (iv) decision-making focus, (v) goals, and (vi) decision-making behavior. An explanation of these fields is provided next.

The *experience* field describes the time in years that the subject has been working in the production and planning department. This time provides a measure of the subject's level of experience.

The *expertise* level is an estimate of the skills that the subject has developed for solving production planning problems. The expertise of participants is expressed in relative terms with regard to the expert decision maker (subject #1) whose expertise level is assumed to be 100%. Other subjects will receive a lower percentage, e.g., 60%, indicating a lower expertise level.

The *role* of a subject in the production planning department describes the main activity of the subject; a subject can be engaged in one of four activities: planning, scheduling, materials management, and customer service.

The *focus* field reflects the relevance that environmental factors have for a subject in his daily activities. A measure of the focus is represented graphically using a line of dots. A five-dot line indicates that subject has a high interest in the corresponding factor; a one-dot line represents a small interest on that factor. For instance, the subject #1 has a high interest in Material Expedited Costs; therefore, five dots (•••••) are marked in his profile for that factor.

The *goals* field indicates the main objectives of the subject in the production planning decisions. A subject may pursue different goals at a different level. A graphical representation using dots is used to represent the goals and the level for each subject.

Finally, the *behavior* is a subjective assessment of the way the subject executes his job. Four categories were used to rate this field. The rating process for the subjects operation's skills uses the same graphical scale (dots). Details are presented in Table 27.

Table 27: Profile of subjects participating in empirical evaluation

Category	**Subject**			
	#1	**#2**	**#3**	**#4**
Experience	15 years	8 years	1 year	12 years
Expertise	100%	70%	30%	60%
Role	Planning	Scheduling	Materials	Customer
Focus				
Material Delivery Improvement	▪▪▪▪	▪▪▪▪▪	▪▪	
Material Expedited Costs	▪▪▪▪▪	▪▪	▪▪▪	
Material Availability Trends	▪▪▪▪▪	▪▪▪▪▪	▪▪	
Competition Price Level	▪▪▪▪▪			
Competition Lead Time	▪▪▪▪▪			
Customer Needed Lead Time		▪▪▪▪▪	▪▪▪	▪▪▪▪▪
Customer Reliability	▪▪▪▪▪	▪▪▪▪	▪▪▪▪▪	▪▪▪▪▪
Market Trends	▪▪▪▪▪	▪▪▪▪		▪▪▪▪▪
Management Production Goals	▪▪▪▪▪	▪▪▪▪▪		
Management Risk Behavior	▪▪▪▪▪			
Unexpected Production Events		▪▪▪▪▪		
Goals				
Meet production goals	▪▪▪▪▪	▪▪▪▪▪	▪▪	
Balance production resources	▪▪▪▪▪	▪▪▪▪▪		
Minimize costs	▪▪▪▪▪		▪▪▪	
Maximize productivity	▪▪▪▪▪			
Customer satisfaction	▪▪▪			▪▪▪▪▪
Behavior				
Detailed analysis	▪▪▪▪▪	▪▪▪	▪▪	▪▪▪
Computer skills	▪▪	▪▪▪▪▪	▪▪▪▪▪	▪
Plays w/diff scenarios	▪▪▪▪▪	▪▪▪		▪
Capable of handling data overload	▪▪▪▪	▪▪▪		▪

APPENDIX D.

Exercises used for the empirical evaluation

Exercise C1

Unit: **Manufacturing Planning**

Topic: **Rough-cut Material Requirements Planning (RC-MRP)**

Suggested topic time: **90 minutes**

Business requirements:

The project management team has determined needs for:

- The creation of the rough-cut MRP for the period from November, 2005 through April, 2006.
- The creation of the inventory replenishment plan to place purchase orders to raw material vendors for the referred planning horizon.
- The participant must assume that the current date is August 1st, 2005.

Business scenario:

- The market is considered to grow gradually for the following 12 months.
- The senior management has authorized increments on inventory levels by of up to 10%.
- The minimum and maximum plant workload has been set to 85% and 95%, respectively.
- The marketing personnel have announced that the levels of the competition price and lead time are higher; they have stated that chances to increase market participation by

5% if the standard delivery times for finished products are improved by 20%. Moreover, they have confirmed that the customers' required delivery times are flexible and they can be negotiated. In this regard, the senior managers are willing to take chances to increase market share.

- The availability of raw materials is showing an increasing trend, i.e., shorter delivery times can be expected from raw material vendors. The purchase department has informed that the raw materials suppliers have offered discounts for expediting material delivery.
- The importance ranking for each external factor has been set by the project management team. A summary of business scenario and importance assessment policies is shown in Table 28.

Required activities:

- The participant must retrieve the required documents from data sources (see Appendix E for a list of available data sources, as well as required transactional commands, to obtain required documents). Alternatively, the participant can access directly the spreadsheets with the data sources. The links to data directory have been created in the "Favorites' node of the Main Access tree (use *miniERP* transactional interface).
- Analyze data
    - Data analysis should include considerations related to business scenario.
    - Pivot tables and charts have been created in each data file.

Deliverables:

- As a result of business analysis, the participant needs to complete the RC-MRP forms. Two types of entries are required:
    - The quantities of each product to manufacture during the planning horizon.
    - The raw material vendor for each product. Associated to each vendor there is a quality index and a price discount (see also Appendix E for more details).
    - A link to the spreadsheet containing the required output is accessible through the 'Favorites' node of the Main Access tree (use ***miniERP*** transactional interface).

Table 28: Environmental conditions for exercise *C1*

External Factors		Importance	State	Business Scenario		
		I	V	Adverse	Neutral	Favorable
mtr01	Material Delivery Improvement	3	3			Immediate
mtr02	Material Expedited Costs	3	3			Low
mtr03	Material Availability Trends	1	3			Abundant
cmp01	Competition Price Level	4	3			Above
cmp02	Competition Lead Time	3	3			Above
mkt01	Customer Needed Lead Time	4	3			Flexible
mkt02	Customer Reliability	2	3			High
mkt03	Market Trends	1	3			Up
dmk01	Acceptable inventory levels	1	3			Up to 10%
dmk02	Management Risk Behavior	1	3			Prone
rdm01	Plant workload levels	4	3			85% - 95%

<u>Considerations and final notes:</u>

- Timing
    - There are not time limitations to deliver required output; however, time is a variable considered in the overall rating.
- Supporting materials are available in Appendix E for participants' consultation. These materials include:

- RCMRP definitions
- RCMRP problem assessment
- A list of available data sources, as well as required transactional commands and paths to navigate the system to reach each transactional window.

Exercise C2

Unit: **Manufacturing Planning**

Topic: **Rough-cut Material Requirements Planning (RC-MRP)**

Suggested topic time: **90 minutes**

<u>Business requirements:</u>

The project management team has determined needs for:

- The creation of the rough-cut MRP for the period from November, 2005 through April, 2006.
- The creation of the inventory replenishment plan to place purchase orders to raw material vendors for the referred planning horizon.
- The participant must assume that the current date is August 1st, 2005.

<u>Business scenario:</u>

- The market is considered to slow down for the following 12 months.
- The senior management has set the policy for zero inventory levels (0%).
- Because of slow business, some layouts are expected; consequently, the maximum plant workload has been set to 80% of installed capacity.

- Marketing personnel have announced that competition has a stronger market position, i.e., price levels and delivery times are better.
- Because of market constraints, the customers' required delivery times are short and not negotiable. Senior management has confirmed their adversity to risky businesses.
- The availability of raw materials is showing a decreasing trend, i.e., longer delivery times can be expected from raw material vendors. The purchase department has informed that costs for expediting raw material delivery are high.
- The importance ranking for each external factor has been set by the project management team. A summary of the business scenario and importance assessment policies is shown in Table 29.

Table 29: Environmental conditions for exercise *C2*

External Factors		Importance	State	Business Scenario		
		I	V	Adverse	Neutral	Favorable
mtr01	Material Delivery Improvement	1	1	NoImprove		
mtr02	Material Expedited Costs	2	1	High		
mtr03	Material Availability Trends	2	1	Scarce		
cmp01	Competition Price Level	3	1	Below		
cmp02	Competition Lead Time	3	1	Below		
mkt01	Customer Needed Lead Time	1	1	Tight & short		
mkt02	Customer Reliability	1	1	Low		
mkt03	Market Trends	1	1	Down		
dmk01	Acceptable inventory levels	2	1	0%		
dmk02	Management Risk Behavior	1	1	Averse		
rdm01	Plant workload levels	2	1	80%		

<u>Required activities:</u>

- Participant must retrieve required documents from data sources (see Appendix E for a list of available data sources, as well as required transactional commands, to obtain required documents). Alternatively, participant can access directly the spreadsheets with the data sources. The links to data directory have been created in the 'Favorites' node of the Main Access tree (use *miniERP* transactional interface).

- Analyze data
    - Data analysis should include considerations related to business scenario.
    - Pivot tables and charts have been created in each data file.

Deliverables:

- As a result of business analysis, the participant needs to complete the RC-MRP forms. Two types of entries are required:
    - The quantities of each product to manufacture during the planning horizon.
    - The raw material vendor for each product. Associated to each vendor there is a quality index and a price discount (see also Appendix E for more details).
    - A link to the spreadsheet containing the required output is accessible through the 'Favorites' node of the Main Access tree (use ***miniERP*** transactional interface).

Considerations and final notes:

- Timing
    - There are not time limitations to deliver required output; however, time is a variable considered in the overall rating.
- Supporting materials are available in Appendix E for participants' consultation. These materials include:
    - RCMRP definitions and problem assessment.

- A list of available data sources, as well as required transactional commands and paths to navigate the system to reach each transactional window.

Exercise C3

Unit: **Manufacturing Planning**

Topic: **Rough-cut Material Requirements Planning (RC-MRP)**

Suggested topic time: **90 minutes**

<u>Business requirements:</u>

The project management team has determined needs for:

- The creation of the rough-cut MRP for the period from November, 2005 through April, 2006.
- The creation of the inventory replenishment plan to place purchase orders to raw material vendors for the referred planning horizon.
- The participant must assume that the current date is August 1st, 2005.

<u>Business scenario:</u>

- The market conditions will show an increasing tendency for the following 12 months.
- Senior management has set the policy for inventory levels to a maximum of 5%.
- Because of business is coming to normal levels, acceptable plant workload has been set to 90% ~ 100%.
- Marketing personnel have announced that competition is offering similar price levels; however, their delivery times are shorter. Senior management is willing to accept some risks in order to maintain and/or gain market share.

- Sales personnel has confirmed that current market share will maintain (old customers' reliability is high). They also have confirmed that old and new customers' required delivery times are flexible and negotiable.
- The raw materials availability is showing an increasing trend. Suppliers can offer immediate delivery time; however, expediting costs are high.
- The importance ranking for each external factor has been set by project management team. A summary of the business scenario and importance assessment policies is shown in Table 30.

Table 30: Environmental conditions for exercise *C3*

External Factors		Importance	State	Business Scenario		
		I	V	Adverse	Neutral	Favorable
mtr01	Material Delivery Improvement	1	1	No improve		
mtr02	Material Expedited Costs	2	1	High		
mtr03	Material Availability Trends	2	1	Scarce		
cmp01	Competition Price Level	3	2		Similar	
cmp02	Competition Lead Time	3	1	Below		
mkt01	Customer Needed Lead Time	1	2		Negotiable	
mkt02	Customer Reliability	1	2		Acceptable	
mkt03	Market Trends	1	3			Increasing
dmk01	Acceptable inventory levels	2	2		Up to 5%	
dmk02	Management Risk Behavior	1	3			Prone
rdm01	Plant workload levels	2	3			90% - 100%

<u>Required activities:</u>

- The participant must retrieve required documents from data sources (see Appendix E for a list of available data sources, as well as required transactional commands, to obtain required documents). Alternatively, participant can access directly the spreadsheets with the data sources. Links to data directory have been created in the 'Favorites' node of the Main Access tree (use *miniERP* transactional interface).
- Analyze data

- Data analysis should include considerations related to business scenario.
- Pivot tables and charts have been created in each data file.

<u>Deliverables:</u>

- As a result of business analysis, the participant needs to complete the RC-MRP forms. Two types of entries are required:
  - The quantities of each product to manufacture during the planning horizon (see Appendix E for further details).
  - The raw material vendor for each product. Associated to each vendor there is a quality index and a price discount (see also Appendix E for more details).
  - A link to the spreadsheet containing the required output is accessible through the 'Favorites' node of the Main Access tree (use ***miniERP*** transactional interface).

<u>Considerations and final notes:</u>

- Timing
  - There are not time limitations to deliver required output; however, time is a variable considered in the overall rating
  - Supporting materials are available in Appendix E for participants' consultation.

APPENDIX E.

Supplementary material for the empirical evaluation

Definitions

<u>Rough-cut Material Requirements Plan (RC-MRP)</u>

The RC-MRP is a schedule of products that the company is planning to produce. It covers a planning horizon of typically 6 periods (months). In this schedule, the production planner inputs the amount of products he thinks the company will sell/manufacture for the planning horizon. Table 31 depicts a draft of the plan.

Table 31: Rough-Cut Material Requirements Plan

Product code	Product description	Quantities (units)					
		Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
BV-2	Ball Valve 2" T31	45	45	55	65	50	50
BV-3	Ball Valve 3" T31	25	25	35	30	30	25
BV-4	Ball Valve 4" T31	30	20	35	35	30	15
BV-6	Ball Valve 6" T31	15	25	50	50	30	30
***							
BV-48	Ball Valve 48" T31	5	3	0	0	2	3

The RC-MRP drives the purchase of raw materials, i.e., it defines the needs of raw materials for the planning horizon. It is called "rough-cut" because it actually doesn't generate requirements for all the materials needed to build the products, but only those that qualify as 'key'. 'Key' raw materials play an important role in the production process due to the long lead time offered by raw material vendors. The importance of an accurate RC-MRP reflects shorter delivery times for future customers' orders. It also impacts the stock levels of raw materials.

Inventory Replenishment Plan (IRP)

The IRP is a schedule of anticipated purchase orders to raw material suppliers for a planning horizon of typically 6 months. In this schedule, the production planner inputs the amount of raw materials that the company plans to acquire from different raw material vendors. A schedule of the raw materials to acquire during the planning horizon is depicted in Table 32. RC-MRP drives the amount of required key raw materials for a horizon planning; however, it doesn't say anything about minimum acquisition costs or minimum delivery times. The IRP reflects this information.

Table 32: RC-MRP and Inventory Replenishment Plan

Product code	Raw material description	Supplier (code) | Quantities (units)											
		Period 1		Period 2		Period 3		Period 4		Period 5		Period 6	
		Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty	Supplier	Qtty
BV-2	BV-2-Seats	MSup-13170	48	MSup-45881	48	MSup-45507	48	MSup-13170	48	MSup-13170	48	MSup-13170	48
BV-2	BV-2-Ball	MSup-46961	50	MSup-45507	50	MSup-45881	50	MSup-45881	50	MSup-46961	50	MSup-46961	50
BV-2	BV-2-Body	MSup-18692	45	MSup-45507	45	MSup-45507	45	MSup-45507	45	MSup-16393	45	MSup-45507	45
BV-2	BV-2-Ends	MSup-43544	40	MSup-10409	40	MSup-16393	40	MSup-45507	40	MSup-43544	40	MSup-45881	40
BV-3	BV-3-Seats	MSup-45881	30	MSup-13170	30	MSup-43544	30	MSup-45507	30	MSup-45881	30	MSup-45881	30
BV-3	BV-3-Ball	MSup-45507	25	MSup-46961	25	MSup-45881	25	MSup-10409	25	MSup-45881	25	MSup-45881	25
BV-3	BV-3-Body	MSup-10409	20	MSup-43544	20	MSup-45507	20	MSup-46961	20	MSup-45881	20	MSup-45507	20
BV-3	BV-3-Ends	MSup-16393	15	MSup-43544	15	MSup-10409	15	MSup-45881	15	MSup-45507	15	MSup-45507	15
***													
BV-48	BV-48-Seats	MSup-45881	15	MSup-45881	15	MSup-45881	15	MSup-13170	15	MSup-46961	15	MSup-45881	15
BV-48	BV-48-Ball	MSup-45507	35	MSup-45881	35	MSup-45881	35	MSup-46961	35	MSup-16393	35	MSup-45507	35
BV-48	BV-48-Body	MSup-45881	50	MSup-45507	50	MSup-45881	30	MSup-45881	30	MSup-16393	30	MSup-10409	0
BV-48	BV-48-Ends	MSup-45507	25	MSup-45507	25	MSup-45507	15	MSup-46961	15	MSup-46961	15	MSup-45507	0

The RC-MRP determines the required amount of production resources, e.g., raw materials. Since an important restriction of the MTO production environment is the minimum inventory levels, it is important to define accurately a schedule for the materials supply. This schedule should define a supplier for each raw material. The limited amount of raw material suppliers, raw materials, costs variation, and differences in raw material quality and delivery time, requires a careful analysis in order to define the best material supplier for each ‘key’ raw material component. The ‘best’ supplier is a function of the cost, delivery time, and quality offered by each.

The importance of an accurate IRP reflects lower costs, shorter delivery times, and better quality of raw materials. The latter has a relevant impact in the production time; the higher the quality the less amount of work is required during the production process.

Problem assessment

Goals:

- Minimize costs.
- Maximize demand satisfaction.
- Minimize decision making time.

Rating criteria:

- Incurred costs: raw material acquisition, inventory holding, manufacturing costs.
- Maximize demand satisfaction.
- Minimize decision making time.

Decisions:

- What raw materials to purchase?
- What quantities of each raw material to purchase?
- When to place purchase orders of raw materials?
- Who to place purchase orders of raw materials?

Information needs[9]:

- Demand requirements by product, by customer, by region.
- Market trends by region and by product.
- Market penetration.
- Probability of receiving orders from customer of different geographical locations.
- Factors that determine and increment probabilities of realizing potential customer orders: price improvement, delivery time improvement, customer-sales relationship.
- Production capacities, planned vs. achieved comparisons.
- Production goals (current and planned).
- Most frequently causes of production time delays.
- Inventory levels by product.
- Raw material supply lead time (achieved vs planned).
- Raw material holding costs.
- Customer-product opportunity costs.
- Competition production-delivery-costs stand.

Knowledge needs:

- Probabilities of realizing any particular potential order.
- Negative effects and their probabilities assessment of adding a potential customer order to the aggregate production plan, e.g., failed assessment, cancellations, increments, lack of financial support.
- Positive effects and their probabilities assessment of adding a potential order.

[9] For the decision in hand these are the pieces of information a decision maker is most interested in. Unfortunately, data extracted from data repositories (Documents – see Appendix C) do not deliver this

Data sources, accessible commands, and menu path

- Data sources[10]:
  - Demand Forecast Report (estimates of customers' demand needs)
    - Command: **MCLH23**
    - Navigation path: **Logistics > Production > SOP > Demand forecast > MCLH23**
  - Production history (production history)
    - Command: **ZM12**
    - Navigation path:
  - Production Backlog (current production commitments)
    - Command: **MDLH23**
    - Navigation path: **Logistics > Production > Production control> Control > Current orders > Display | MDLH23**
  - Inventory control report (current inventory levels status)
    - Command: **MD04**
    - Navigation path#1: **Logistics > Production > Production planning> Demand management > Environment > Stock / requirements List | MD04**
    - Navigation path#2: **Logistics > Production > Production control> Control > Stock / requirements List | MD04**
  - Raw material suppliers contracts (description of raw material suppliers analysis)

---

information directly.

[10] For the decision in hand, these are the data sources decision maker uses most frequently. According to the NB-DSS terminology, these are elements are called "Documents" data

        - Command: **MSLH123, MSLH153**
        - Navigation path#1: **Logistics** > **Production** > **SOP** > **Suppliers** > **Display | MSLH123**
        - Navigation path#2: **Logistics** > **Production** > **Master data** > **Production resources and tools** > **Material suppliers** > **Display | MSLH153**
    - Products catalog (description of products and standard configuration options)
        - Command: **MM03**
        - Navigation path: **Logistics** > **Production** > **Master data** > **Material** > **Display | MM03**
    - Products' bill of materials data (description of raw material needs for each product type)
        - Command: **CS03**
        - Navigation path: **Logistics** > **Production** > **Master data** > **Bills of material** > **Bill of material** > **Material BOM** > **Display | CS03**
    - Products' routing data (description of production resources needs and manufacturing times for each product type)
        - Command: **CR03**
        - Navigation path: **Logistics** > **Production** > **Master data** > **Routings** > **Routings** > **Standard routings** > **Display | CR03**

- Output forms
    - Rough-cut MRP
        - Command: **MCLH21**

    - Navigation path: **Logistics > Production > SOP > SOP Demand forecasting > Create | MCLH21**
- Inventory replenishment plan (IRP)
    - Command: **MILH121**
    - Navigation path: Logistics > Production > SOP > SOP Inventory Planning > Create | MILH121

APPENDIX F.

Sample output of transactions during a problem solving session

The sample form presented in **Table 33** corresponds to the transactions executed for the following treatment:

- Interface type: “Control Condition C, using the *miniERP* interface”
- Subject: “Expert”
- Problem difficulty level: “Favorable”

Description of fields

The following fields are used in the sample form:

- “Transaction time”: The start time of each transaction in dd/mm/yyyy, hh:mm:ss format
- “Document / transaction type”: Two types of values are possible; (OFF) when the current transaction is dedicated to offline analysis, i.e., to use the content of current document for framing the problem at hand, (not-OFF) when the focus of current transaction is to access or creation of a document. Depending on the type of generic document, there can exist seven types of summary reports: DF-Demand Forecast, PB-Production Backlog, PH-Production History, SC-Supply Contracts, VD-Vendor Data, PC-Production Capacity, and CO-Product Configuration options.
- “Current Document Num”: This is a sequential number referring to the number of documents that decision maker has created for his support during the problem-solving process. When a new document (‘summary report’) is created this list increments by

one. The same number may appear in subsequent transactions indicating that current document is being used for offline analysis.

- "Current Document has codification”: This is a binary variable that is applicable every time a new document (‘summary report’) has been created. A ‘yes’ value indicates that current document exists in the list of documents contained in the “Expert knowledge”. Recall that the list of document contained in the “Expert knowledge” was built during the ethnographic study; this list comprises all the documents that were recognized as relevant or used during the ethnographic.
- “Transaction description”: A text description of the type of transaction being executed.

Table 33: Sample output of transactions during a problem-solving session

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 13:45	OFF	0		Logged to system
3/29/2007 13:45:30	OFF	0		Execute 'mclh23' transaction
3/29/2007 13:46:15	DF	1	Yes	Accessed Demand Forecast Data
3/29/2007 13:47:00	OFF	1		Export Demand Forecast Report to spreadsheet
3/29/2007 13:47:15	OFF	1		Execute 'mdlh23' transaction
3/29/2007 13:48:00	PB	2	Yes	Accessed Production Backlog Data
3/29/2007 13:48:45	OFF	2		Export Production Backlog Report to spreadsheet
3/29/2007 13:49:00	OFF	2		Execute 'zm14' transaction
3/29/2007 13:49:45	PH	3	Yes	Accessed Production History Data
3/29/2007 13:50:30	OFF	3		Export Production History Report to spreadsheet
3/29/2007 13:50:45	OFF	3		Excute 'mslh123' transaction
3/29/2007 13:51:30	SC	4	Yes	Accessed Supply Contracts Data
3/29/2007 13:52:15	OFF	4		Export Supply Contracts Report to spreadsheet
3/29/2007 13:52:30	OFF	4		Execute 'zr485' transaction
3/29/2007 13:53:15	PC	5	Yes	Accessed Production Capacity Data
3/29/2007 13:54:00	OFF	5		Export Production Capacity Report to spreadsheet
3/29/2007 13:54:15	OFF	5		Open Jexcel application
3/29/2007 13:54:30	DF	6	Yes	Open **Demand Forecast** Report, generic report
3/29/2007 13:54:45	OFF	6		Create Demand Forecast Report, pivot report
3/29/2007 13:59:45	DF	7		Create Demand Forecast Report, pivot chart - show (A)Demand (B)BookDate_All (C)None
3/29/2007 14:00:45	DF	8	Yes	Create Demand Forecast Report, pivot chart - show (A)Demand (B)BookDate_All (C)Line_All
3/29/2007 14:01:45	DF	9	Yes	Create Demand Forecast Report, pivot chart - show (A)Demand (B)BookDate_All (C)Prob_All
3/29/2007 14:02:45	DF	10		Create Demand Forecast Report, pivot chart - show (A)Demand (B)BookDate_All, Margin_345 (C)Line_All
3/29/2007 14:03:45	DF	11	Yes	Create Demand Forecast Report, pivot chart - show(A)Demand (B)BookDate_All (C)Region_All
3/29/2007 14:04:45	DF	12		Create Demand Forecast Report, pivot chart - show (A)Demand (B)Line_All (C)None
3/29/2007 14:05:45	OFF	12		!Conclude demand / market is growing
3/29/2007 14:06:30	PH	13	Yes	Open **Production History** Report, generic report
3/29/2007 14:06:45	OFF	13		Create Production History Report, pivot report
3/29/2007 14:11:45	PH	14		Report of a 13-month period for the previous year, all sizes
3/29/2007 14:12:45	PH	15		Report of a 13-month period for the previous year, product line 2"~6"
3/29/2007 14:13:45	PH	16	Yes	Report of a 13-month period for the previous year, product line 2"~6", add details by size
3/29/2007 14:14:45	OFF	16		!Concludes market has maintained steady (flat behavior). Compute average monthly production for the 2"~6" product line
3/29/2007 14:15:30	OFF	16		# Summarized quantities produced by month and obtained average monthly production 479 pc
3/29/2007 14:17:30	OFF	16		# Summarized quantities produced by month and obtained average weekly production 110 pc
3/29/2007 14:19:30	PC	17	Yes	Open **Production Capacity** Report, generic report
3/29/2007 14:19:45	OFF	17		Create Production Capacity Report, pivot report
3/29/2007 14:24:45	PC	18		Stated production capacity, all lines, all grades
3/29/2007 14:25:15	PC	19		Stated production capacity, all lines, all sizes, all grades
3/29/2007 14:25:43	PC	20		Stated production capacity, all lines, grades A&B
3/29/2007 14:26:16	PC	21	Yes	Stated production capacity, lines: 2~6", grades A&B
3/29/2007 14:27:02	OFF	21		# Computed quantities: 163 pc/week with quality A vs 108 pc/week with quality B
3/29/2007 14:30:02	OFF	21		! Conclude there are opportunities to increase production if acquire only raw material of quality A
3/29/2007 14:31:02	OFF	21		# Compute increase production factor for line 2~6": 163/110 = 1.48 (48%) extra if acquire only raw material of quality A

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 14:32:32	SC	22	Yes	Open **Supply Contracts** Report, generic report
3/29/2007 14:32:47	OFF	22		Create Supply Contracts Report, pivot report
3/29/2007 14:36:47	SC	23		Report of raw material suppliers for 2~6"
3/29/2007 14:37:17	SC	24		Report of raw material suppliers for 2~6", quality A&B
3/29/2007 14:37:53	SC	25		Report of raw material suppliers for 2~6", quality A&B, sort by discount level
3/29/2007 14:38:20	SC	26	Yes	Report of raw material suppliers for 2~6", quality A&B, show delivery time
3/29/2007 14:38:46	SC	27		Report of raw material suppliers for 2~6", quality A&B, show minimum batch size
3/29/2007 14:39:26	SC	28	Yes	Report of raw material suppliers for 2~6", quality A&B, sort by quality and discount
3/29/2007 14:40:12	OFF	28		# Compute best delivery time for quality A raw material suppliers, pick supplier MSup-43573
3/29/2007 14:43:12	OFF	28		Execute 'milh121' transaction
3/29/2007 14:43:27	OFF	28		Accessed RC-MRP output form
3/29/2007 14:43:57	OFF	28		Export RC-MRP form to spreadsheet
3/29/2007 14:44:12	OFF	28		Open RC-MRP output form
3/29/2007 14:44:18	OFF	28		Define RC-MRP for product size 2"
3/29/2007 14:44:24	OFF	28		#Compute average historic monthly production for 2" product size: 83 pc
3/29/2007 14:45:39	PH	29	Yes	Report of a 13-month period for the previous year, product line 2"~6", add details by size=2
3/29/2007 14:46:06	OFF	29		#Compute maximum increasing factor for 2"~6" product line: 48%
3/29/2007 14:46:40	PC	30	Yes	Stated production capacity, lines: 2~6", size =2, grades A&B
3/29/2007 14:47:25	OFF	30		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 14:47:40	OFF	30		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 14:48:40	OFF	30		#Compute average monthly production for planning horizon: (83)(1.48)(0.85) = 103 pc Nov (distributed evenly for 4 End types)
3/29/2007 14:49:25	OFF	30		#Compute target production level for each end conn type: 103/4=> 25, 26, 27, 25 (150#, 300#, 600#, We)
3/29/2007 14:50:25	OFF	30		Open RC-MRP output form
3/29/2007 14:50:31	OFF	30		!Enter values for product size 2", for the entire planning horizon
3/29/2007 14:51:31	OFF	30		Open RC-MRP output form
3/29/2007 14:51:37	OFF	30		Define RC-MRP for current product size: **3"**, current product line: **2"~6"**
3/29/2007 14:51:43	OFF	30		#Compute average historic monthly production for current product size: X pc
3/29/2007 14:52:58	PH	31	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 14:53:25	OFF	31		#Compute maximum ***increasing factor*** for current product line: 48%
3/29/2007 14:54:00	PC	32	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 14:54:45	OFF	32		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 14:55:00	OFF	32		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 14:56:00	OFF	32		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 14:56:45	OFF	32		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 14:57:45	OFF	32		Open RC-MRP output form
3/29/2007 14:57:51	OFF	32		!Enter values for current product size, for the planning horizon
3/29/2007 14:58:51	OFF	32		Open RC-MRP output form
3/29/2007 14:58:57	OFF	32		Define RC-MRP for current product size: **4"**, current product line: **2"~6"**
3/29/2007 14:59:03	OFF	32		#Compute average historic monthly production for current product size: X pc
3/29/2007 15:00:18	PH	33	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 15:00:45	OFF	33		#Compute maximum ***increasing factor*** for current product line: 48%
3/29/2007 15:01:19	PC	34	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 15:02:04	OFF	34		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 15:02:19	OFF	34		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 15:03:19	OFF	34		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 15:04:04	OFF	34		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 15:05:04	OFF	34		Open RC-MRP output form
3/29/2007 15:05:10	OFF	34		!Enter values for current product size, for the planning horizon
3/29/2007 15:06:10	OFF	34		Open RC-MRP output form
3/29/2007 15:06:16	OFF	34		Define RC-MRP for current product size: **6"**, current product line: **2"~6"**
3/29/2007 15:06:22	OFF	34		#Compute average historic monthly production for current product size: X pc
3/29/2007 15:07:37	PH	35	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 15:08:04	OFF	35		#Compute maximum ***increasing factor*** for current product line: 48%
3/29/2007 15:08:38	PC	36	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 15:09:23	OFF	36		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 15:09:38	OFF	36		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 15:10:38	OFF	36		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 15:11:23	OFF	36		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 15:12:23	OFF	36		Open RC-MRP output form
3/29/2007 15:12:29	OFF	36		!Enter values for current product size, for the planning horizon
3/29/2007 15:13:29	OFF	36		Access **Production History** Report, pivot report
3/29/2007 15:13:44	PH	37		Report of a 13-month period for the previous year, all sizes
3/29/2007 15:14:14	PH	38		Report of a 13-month period for the previous year, product line 8"~12"
3/29/2007 15:14:44	PH	39	Yes	Report of a 13-month period for the previous year, product line 8"~12", add details by size
3/29/2007 15:15:14	OFF	39		!Identify peaks and valleys. Compute average monthly production for the 8"~12" product line. Remove peaks and valleys

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 15:17:14	OFF	39		# Summarized quantities produced by month and obtained average monthly production 99 pc
3/29/2007 15:18:14	OFF	39		# Summarized quantities produced by month and obtained average weekly production 22 pc
3/29/2007 15:19:14	OFF	39		Access **Production Capacity** Report, pivot report
3/29/2007 15:19:29	PC	40		Stated production capacity, all lines, all grades
3/29/2007 15:19:59	PC	41		Stated production capacity, all lines, all sizes, all grades
3/29/2007 15:20:27	PC	42		Stated production capacity, all lines, grades A&B
3/29/2007 15:21:00	PC	43	Yes	Stated production capacity, lines: 8~12", grades A&B
3/29/2007 15:21:46	OFF	43		# Computed quantities: 38 pc/week with quality A vs 21 pc/week with quality B
3/29/2007 15:22:46	OFF	43		! Conclude there are opportunities to increase production if acquire only raw material of quality A
3/29/2007 15:23:16	OFF	43		# Compute increase production factor for raw material of quality A
3/29/2007 15:23:46	OFF	43		Access **Supply Contracts** Report, pivot report
3/29/2007 15:24:01	SC	44		Report of raw material suppliers for 8~12"
3/29/2007 15:24:31	SC	45		Report of raw material suppliers for 8~12", quality A&B
3/29/2007 15:25:07	SC	46		Report of raw material suppliers for 8~12", quality A&B, sort by discount level
3/29/2007 15:25:34	SC	47	Yes	Report of raw material suppliers for 8~12", quality A&B, show delivery time
3/29/2007 15:26:00	SC	48		Report of raw material suppliers for 8~12", quality A&B, show minimum batch size
3/29/2007 15:26:40	SC	49	Yes	Report of raw material suppliers for 8~12", quality A&B, sort by quality and discount
3/29/2007 15:27:26	OFF	49		# Compute best delivery time for quality A raw material suppliers. Select supplier
3/29/2007 15:28:26	OFF	49		Open RC-MRP output form
3/29/2007 15:28:32	OFF	49		Define RC-MRP for current product size: **8"**, current product line: **8"~12"**
3/29/2007 15:28:38	OFF	49		#Compute average historic monthly production for current product size: X pc
3/29/2007 15:29:53	PH	50	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 15:30:20	OFF	50		#Compute maximum ***increasing factor*** for current product line: **72%**
3/29/2007 15:30:55	PC	51	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 15:31:40	OFF	51		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 15:31:55	OFF	51		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 15:32:55	OFF	51		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 15:33:40	OFF	51		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 15:34:40	OFF	51		Open RC-MRP output form
3/29/2007 15:34:46	OFF	51		!Enter values for current product size, for the planning horizon
3/29/2007 15:35:46	OFF	51		Open RC-MRP output form
3/29/2007 15:35:52	OFF	51		Define RC-MRP for current product size: **10"**, current product line: **8"~12"**
3/29/2007 15:35:58	OFF	51		#Compute average historic monthly production for current product size: X pc

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 15:37:13	PH	52	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 15:37:40	OFF	52		#Compute maximum ***increasing factor*** for current product line: **72%**
3/29/2007 15:38:14	PC	53	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 15:38:59	OFF	53		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 15:39:14	OFF	53		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 15:40:14	OFF	53		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 15:40:59	OFF	53		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 15:41:59	OFF	53		Open RC-MRP output form
3/29/2007 15:42:05	OFF	53		!Enter values for current product size, for the planning horizon
3/29/2007 15:43:05	OFF	53		Access **Production History** Report, pivot report
3/29/2007 15:43:11	PH	54		Report of a 13-month period for the previous year, all sizes
3/29/2007 15:43:41	PH	55		Report of a 13-month period for the previous year, product line **14" - larger**
3/29/2007 15:44:11	PH	56	Yes	Report of a 13-month period for the previous year, product line **14" - larger**, add details by size
3/29/2007 15:44:41	OFF	56		!Identify and remove peaks and valleys for ea product size. Compute average monthly production for the 14"- larger product line.
3/29/2007 15:45:41	OFF	56		# Summarized quantities produced by month and obtained average monthly production X pc
3/29/2007 15:46:11	OFF	56		# Summarized quantities produced by month and obtained average weekly production Y pc
3/29/2007 15:46:41	OFF	56		Access **Production Capacity** Report, pivot report
3/29/2007 15:46:47	PC	57		Stated production capacity, all lines, all grades
3/29/2007 15:47:17	PC	58		Stated production capacity, all lines, all sizes, all grades
3/29/2007 15:47:44	PC	59		Stated production capacity, all lines, grades A&B
3/29/2007 15:48:18	PC	60	Yes	Stated production capacity, lines: **14" - larger**, grades A&B
3/29/2007 15:49:04	OFF	60		# Computed quantities: M pc/week with quality A vs N pc/week with quality B
3/29/2007 15:49:34	OFF	60		! Conclude there are opportunities to increase production if acquire only raw material of quality A
3/29/2007 15:49:49	OFF	60		# Compute increase production factor for raw material of quality A
3/29/2007 15:50:04	OFF	60		Access **Supply Contracts** Report, pivot report
3/29/2007 15:50:10	SC	61		Report of raw material suppliers for **14" - larger**
3/29/2007 15:50:40	SC	62		Report of raw material suppliers for **14" - larger**, quality A&B
3/29/2007 15:51:16	SC	63		Report of raw material suppliers for **14" - larger**, quality A&B, sort by discount level
3/29/2007 15:51:43	SC	64	Yes	Report of raw material suppliers for **14" - larger**, quality A&B, show delivery time
3/29/2007 15:52:09	SC	65		Report of raw material suppliers for **14" - larger**, quality A&B, show minimum batch size
3/29/2007 15:52:49	SC	66	Yes	Report of raw material suppliers for **14" - larger**, quality A&B, sort by quality and discount
3/29/2007 15:53:35	OFF	66		# Compute best delivery time for quality A raw material suppliers. Select supplier
3/29/2007 15:53:50	OFF	66		Open RC-MRP output form

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 15:53:56	OFF	66		Define RC-MRP for current product size: **16"**, current product line: **14"-larger**
3/29/2007 15:54:02	OFF	66		#Compute average historic monthly production for current product size: X pc
3/29/2007 15:55:17	PH	67	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 15:55:44	OFF	67		#Compute maximum ***increasing factor*** for current product line: **IF%**
3/29/2007 15:56:18	PC	68	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 15:57:03	OFF	68		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 15:57:18	OFF	68		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 15:58:18	OFF	68		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 15:59:03	OFF	68		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 16:00:03	OFF	68		Open RC-MRP output form
3/29/2007 16:00:09	OFF	68		!Enter values for current product size, for the planning horizon
3/29/2007 16:01:09	OFF	68		Open RC-MRP output form
3/29/2007 16:01:15	OFF	68		Define RC-MRP for current product size: **24"**, current product line: **14"-larger**
3/29/2007 16:01:21	OFF	68		#Compute average historic monthly production for current product size: X pc
3/29/2007 16:02:36	PH	69	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 16:03:03	OFF	69		#Compute maximum ***increasing factor*** for current product line: **IF%**
3/29/2007 16:03:37	PC	70	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 16:04:22	OFF	70		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 16:04:37	OFF	70		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 16:05:37	OFF	70		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 16:06:22	OFF	70		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 16:07:22	OFF	70		Open RC-MRP output form
3/29/2007 16:07:28	OFF	70		!Enter values for current product size, for the planning horizon
3/29/2007 16:08:28	OFF	70		Open RC-MRP output form
3/29/2007 16:08:35	OFF	70		Define RC-MRP for current product size: **36"**, current product line: **14"-larger**
3/29/2007 16:08:41	OFF	70		#Compute average historic monthly production for current product size: X pc
3/29/2007 16:09:56	PH	71	Yes	Report of a 13-month period for the previous year, current product line, add details for current product size
3/29/2007 16:10:23	OFF	71		#Compute maximum ***increasing factor*** for current product line: **IF%**
3/29/2007 16:10:57	PC	72	Yes	Retrieve production capacity report for current production lines, current product size, raw material quality
3/29/2007 16:11:42	OFF	72		!Use fact: maximum plant workload factor: 85%~95%

Table 33 (Continued)

Transaction time	Document | Transaction type	Current Doc Num	Doc has code (Y/N)	Transaction description
3/29/2007 16:11:42	OFF	72		!Use fact: maximum plant workload factor: 85%~95%
3/29/2007 16:11:57	OFF	72		#Distribute increasing factor during planning horizon, 0.85(P1), 0.87(P2), 0.89(P3), 0.91(P4), 0.93(P5), 0.95(P6)
3/29/2007 16:12:57	OFF	72		#Compute average monthly production for planning horizon: (X pc)(1.48)(0.85) = Y pc Nov (distributed evenly for 4 End types)
3/29/2007 16:13:42	OFF	72		#Compute target production level for each end conn type: Y/4=> Y1, Y2, Y3, Y4 (150#, 300#, 600#, We)
3/29/2007 16:14:42	OFF	72		Open RC-MRP output form
3/29/2007 16:14:48	OFF	72		!Enter values for current product size, for the planning horizon
			42	

2:30:48 Total time required to compute a solution:
0:44:40 Total time spent in search of data and documents:
1:46:08 Total time spent on offline analysis:
72 Total number of documents created:
42 Total number of documents created and used:
0 Total number of documents created and used, but w/o codification

APPENDIX G.

Parameters used to compute cost-based performance measure

Holding Costs:

Inventory holding cost for any product was computed at 3% of its cost.

Acquisition costs:

Table 34: Acquisition costs for all products

Product size	Product line	Acquisition Cost
2	02" - 06"	1,000.00
3	02" - 06"	1,250.00
4	02" - 06"	1,400.00
6	02" - 06"	2,100.00
8	08" - 12"	3,200.00
10	08" - 12"	3,800.00
12	08" - 12"	5,200.00
16	14" - larger	7,500.00
20	14" - larger	11,000.00
24	14" - larger	15,000.00
30	14" - larger	21,000.00
36	14" - larger	30,000.00

Costs in US$

Opportunity costs:

Table 35: Gross margin levels for each product line

Product line	Gross margin levels
2"~6"	55%
8"~12"	65%
14"~36"	70%

APPENDIX H.

Empirical evaluation protocol

In this appendix we include two documents: the IRB Application and the Consent – Confidentiality – Risk.

Document 1: IRB Application

I. General Protocol Information

A. **Protocol title:** "Evaluation of the miniERP-NDSS software in manufacturing production planning".

B. **Research Personnel**: Dr. T. Govindaraj - Principal Investigator, Luis E. Herrera – Co-investigator.

C. **Protocol description**:

(See Chapter 6 of this dissertation).

D. **Protocol Department**: ISyE (depending on list of choices).

E. **Research funding**: N/A.

F. **Research locations**: Manufacturing plant located in Ville Platte, LA.

II. Lay Summary

A. **Certification**: Completed and passed the Human Subjects Training requirement.

B. **Describe in lay terms the purpose of the research including the research question**:

Research purpose:

This research is aimed at empirically evaluating a Network-based Decision Support System proof-of-concept implementation (miniERP-NDSS software). The evaluation will assess the performance and effectiveness of the software in supporting production planners during the creation of a rough-cut material requirements plan (RC-MRP).

Research question:

Does the 'miniERP-NDSS' implementation improve the performance of production planners in the achievement of their goals? (*)

(*) Goals include:

- Completion time of planning tasks
- Minimization of inventory levels
- Maximization of demand satisfaction
- Minimization of production time
- Minimization of production costs
- "To do the most with the least in the least time"

What do you hope to gain by doing this research?

We hope to better understand the usefulness and effectiveness of a network-based modeling approach to support decision making. Two main measures of effectiveness are being evaluated:

- The effectiveness of the network-based model and applied visualization techniques to improve the decision making performance.
- The effectiveness of the network-based model approach to capture and structure the domain knowledge (qualitative data) usually 'locked up' in decision makers' heads.

III. Subject Information

A. **Human Subject Interaction**: Yes, there is direct interaction with human subjects.

B. **Proposed Consent Procedures**: Adult consent form, see next document.

C. **HIPAA Questions**: To be answered in the website.

D. **Subjects Data**: To be answered in the website.

E. **Type of review requested**: Exempt.

F. **Will the study involve Drugs**: No.

G. **Will the study involve Investigational Devices:** Yes, a laptop computer.

H. **Will the study involve Radiation:** No.

IV. Other Questions

A. **Does this research activity involve collection of biological specimens?** N/A.

B. **If prospective, specify:** N/A.

C. **Will specimens be collected anonymously?** N/A.

D. **Is genetic testing of these specimens proposed?** N/A.

E. **Has BSC approval been obtained?** N/A.

F. **Is the use of rDNA proposed?** N/A.

G. **Check for documentation you will upload (actual upload is made in section VI):**

a. Consent form – Yes. (uploaded)

b. Grant – N/A.

c. Other committee or institute approval letters – N/A.

d. Proposal / Dissertation – Yes.

e. Recruitment materials – N/A.

f. Surveys / Questionnaires – N/A.

V. Key words that describe this protocol

A. Decision, assessment, goals.

Document #2: Consent – Confidentiality – Risk

Georgia Institute of Technology

Project Title: "Evaluation of the *miniERP-NDSS* Network-based Decision Support System in Manufacturing Production Planning"

Investigators: Dr. T. Govindaraj - Principal Investigator, Luis E. Herrera – Co-Investigator

<u>Research Consent Form</u>

You are being asked to be a volunteer in a research study. Details about the research purpose and methodologies are described next.

<u>Purpose:</u>

The purpose of this study is:

This research pursues an empirical evaluation of the Network-based Decision Support System proof-of-concept implementation (*miniERP-NDSS* Assistant). The evaluation assesses the performance and effectiveness of the assistant in supporting production planners during the creation of a rough-cut material requirements plan (RC-MRP).

<u>Procedures:</u>

Volunteers for this study will participate in two free-time problem solving sessions. Each session is expected to last from 60 to 120 minutes. In the first session participants are asked to execute/solve three regular problems of rough-cut material requirements plan utilizing the provided transactional interface of miniERP-NDSS. Each problem varies on difficulty level and environmental circumstances.

The second session provides follow-on to the three problems presented on the first session. In the second session participants are asked to solve the same set of problems, but utilizing the miniERP-NDSS software assistant.

It is expected that participants will complete one session per day for two consecutive days. Within a session, participants will be given the opportunity for short breaks between each problem. Due to their normal work schedules within which this evaluation study will be embedded, the participants may take longer breaks (more than 60 minutes) or even miss a day between problems corresponding to a single session. Table 36 summarizes the contents of the two sessions together with the approximate duration of each.

Table 36: Overview of RC-MRP problem solving sessions

Session	Purpose	Activities	Duration
1	Test problem solving performance and accuracy with standard interface	• Introduction	3 minutes
		• System description and explanation	10 minutes
		• Lesson A (*): miniERP Transactional interface training	20 minutes
		• Instructions for problem #1	5 minutes
		• Problem #1 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #2	5 minutes
		• Problem #2 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #3	5 minutes
		• Problem #3 – timed RC-MRP execution	Free

Table 36 (continued)

2	Test problem solving performance and accuracy with decision support system interface	• Introduction	3 minutes
		• System description and explanation	10 minutes
		• Lesson B (*): miniERP-NDSS Decision Support Interface training	20 minutes
		• Instructions for problem #1	5 minutes
		• Problem #1 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #2	5 minutes
		• Problem #2 – timed RC-MRP execution	free
		• Break	15 min - free
		• Instructions for problem #3	5 minutes
		• Problem #3 – timed RC-MRP execution	Free

(*) Prior to each problem solving session, participants will receive two training lessons (A and B). These lessons will permit participants to become familiar with the interfaces (transactional and decision support) required to solve the problems. More details on the activities in each session are included in Chapter 6 of this dissertation. All activities in each session will be video taped to better understand their actions.

<u>Risks/Discomforts</u>

The risks involved are no greater than those involved during the normal execution of tasks during their regular work schedule.

<u>Benefits</u>

Participants will be granted access to *minERP-NDSS* interface for continuing evaluation and improvement. They are not likely to benefit in any other way than for training purposes addressing the problems herein treated.

<u>Compensation to You</u>

Participants agree to not receive any monetary compensation for participating in this study.

Confidentiality:

The following procedures will be followed to keep your personal information confidential in this study:

- Collected data about you will be kept private to the extent allowed by law. To protect your privacy, your records will be kept under a code number rather than by name. Your records will be kept in locked files and only study staff will be allowed to look at them. Your name and any other fact that might point to you will not appear when results of this study are presented or published.
- Video and audio tape data obtained from your participation will be used only by the principal investigators (PI) to derive conclusions of this study. After the study is ended and all necessary information has been obtained, PI will keep these data for archival purposes.
- Experiments will be run on the computer (hardware) provided by the Principal Investigators (PI). Data from the experiments, including the data you entered in your responses will be stored and kept by the PI. After analysis of data, these will be kept for archival purposes.

To make sure that this research is being carried out in the proper way, the Georgia Institute of Technology IRB will review study records. Members of the Food and Drug Administration may also look over study records during required reviews. The Office of Human Research Protections may also look at study records.

Costs to You

Your participation on this study will not carry any associated costs.

Alternative Treatments N/A

Subject Rights

As a participant of this study you should be aware of the following:

- Your participation in this study is voluntary.
- You do not have to be in this study if you don't want to be.
- You have the right to change your mind and leave the study at any time without giving any reason, and without penalty.
- Any new information that may make you change your mind about being in this study will be given to you.
- You will be given a copy of this consent form to keep.
- You do not waive any of your legal rights by signing this consent form.

Questions about the Study or Your Rights as a Research Subject

If you have any questions about the study, you may contact Dr. T. Govindaraj, at telephone (404) 894-3873. If you have any questions about your rights as a research subject, you may contact Ms. Alice Basler, Georgia Institute of Technology at (404) 894-6942. If you sign below, it means that you have read (or have had read to you) the information given in this consent form, and you would like to be a volunteer in this study.

____

Subject Name

Subject Signature **Date**

Signature of Person Obtaining Consent **Date**

REFERENCES

Andersen, P., & Petersen, N. C. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science, 39*(10).

Anonymous. (2007a). JUNG - Java Universal Network/Graph Framework. Retrieved July/2007, from http://jung.sourceforge.net/

Anonymous. (2007b). Prefuse Information Visualization Toolkit. Retrieved August/2007, 2007, from http://prefuse.org

Anthony, R. N. (1965). Planning and Control Systems: A Framework for Analysis. *Harvard University Graduate School of Business Administration*.

Bailey, K. D. (1994). *Typologies and taxonomies: an introduction to classification techniques*. Thousand Oaks, Cal.: Sage Publications.

Banker, R. D., Charnes, A., & Cooper, W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science, 30*(9), 1078-1092.

Banker, R. D., & Morey, R. C. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research, 34*(4), 513-521.

Barry, C. A. (1998). Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared. *Sociological Research Online* Retrieved April 16, 2007, from http://www.socresonline.org.uk/

Benamati, J., Lederer, A. L., & Singh, M. (1997). Changing information technology and information technology management. *Information and Management, 31*(5), 275-288.

Bendoly, E. (2003). Theory and support for process frameworks of knowledge discovery and data mining from ERP systems. *Information and Management, 40*(7), 639-647.

Blismas, N. G. (2003). Computer-aided qualitative data analysis: Panacea or paradox? *Building Research and Information, 31*(6), 455-463.

Bonczek, R. H., & Holsapple, A. B. (1981). *Foundations of Decision Support Systems*. New York: Academic Press.

Cassell, C., Anna, B., Gillian, S., & Johnson, P. (2006). Qualitative Methods in Management Research: An Introduction to the Themed Issue. *Management Decision, 44*(2), 161-166.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research 2*(6), 429-444.

Chaudhuri, S. (1997). An overview of data warehousing and OLAP technology. *ACM SIGMOD, 26*(1), 65-74.

Cook, W., & Kress, M. (1990). A Data Envelopment Model for Aggregating Preference Rankings. *Management Science, 36*(11).

Cook, W., & Zhu, J. (2006). Rank order data in DEA: A general framework. *European Journal of Operational Research, 174*(2), 1021-1038.

Courtney, J. F., & Parandise, D. B. (1993). Studies in Managerial Problem Formulation Systems. *Decision Support Systems, 9*(4), 413-423.

Dey, I. (1993). *Qualitative Data Analysis: A User-Friendly Guide* (First ed.). London: Routledge.

Doty, D. H., & Glick, W. H. (1994). Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling *The Academy of Management Review, 19*(2), 230-251.

Drucker, P. (1988). The Coming of New Organization. *Harvard Business Review*(Jan-Feb), 46.

Duncan, R. B. (1972). Characteristics of Organizational Environments and Perceived Environmental Uncertainty. *Administrative Science Quarterly, 17*, 313-327.

Dyson, R. G., & Thanassoulis, E. (1988). Reducing Weight Flexibility in Data Envelopment Analysis. *Journal of Operational Research Society, 39*(6), 563-576.

Ertay, T., Ruan, D., & Tuzcaya, U. r. (2006). Integrating Data Envelopment Analysis and Hierarchy for the Facility Layout Design in Manufacturing Systems. *Information Systems, 176*, 237-262.

Fulkerson, W. (2000). Information-based Manufacturing in the Informational Age. *The International Journal of Flexible Manufacturing Systems, 12*, 131-143.

Golany, B. (1988). A Note on Including Ordinal Relations Among Multipliers in Data Envelopment Analysis. *Management Science, 34*(8), 1029-1033.

Gorry, G. A., & Scott Morton, M. S. (1971). A Framework for Management Information Systems. *Sloan Management Review 13*(1), 50-70.

Groover, M. (2002). *Automation, Production systems, and Computer-Integrated Manufacturing* (Second ed.).

Herrera, L. E., & Komischke, T. (2007, July 22-27). *Coping with Information Input Overload: User Interface Concepts for Industrial Process Control.* Paper presented at the HCII International 2007, Beijing.

Horvitz, E., Azari, D., & Dumais, S. (2004). Actions, answers, and uncertainty: a decision-making perspective on Web-based question answering. *Information Processing & Management, 40*(5), p849-868.

Hu, P. J.-H., Ma, P.-C., & Chau, P. Y. K. (1999). Evaluation of user interface designs for information retrieval systems: a computer-based experiment. *Decision Support Systems*, *27*, 125–143.

Huber, G. P. (1983). Cognitive Style as a Basis for MIS and DSS Design: Much Ado About Nothing? *Management Science*, *29*(5), 567-577.

Lalmas, M., & Ruthven, I. (1998). Representing and retrieving structured documents using the Dempster-Shafer theory of evidence: modeling and evaluation. *Journal of Documentation, 54*(5), 529-565.

Lant, T., K.,, & Hewlin, P. (2002). Information Cues and Decision Making. *Group & Organization Management*, *27(3)*(3), 373-407.

Leavitt, H. J. (1976). Beyond the analytic manager. *California Management Review, 17*(3), 5-12.

Mintzbergh, H., & Raisinghani, D. (1976). The Structure of "Unstructured" Decision Processes. *Administrative Science Quarterly*, *21*(2), 246-275.

Murdock, V., Kelly, D., Croft, W. B., Belkin, N. J., & Yuan, X. (2007). Identifying and improving retrieval for procedural questions. *Information Processing and Management*, *43*(1), 181-203.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review*, *E69*(E69), 026113-026111~026115.

OLAPCouncil. (1997). OLAP Council White Paper. Retrieved February, 2007, from http://www.symcorp.com/downloads/OLAP_CouncilWhitePaper.pdf

Paradi, J. C., Reese, D. N., & Rosen, D. (1995). Applications of DEA to Measure the Efficiency of Software Production at Two Large Canadian Banks. *Annals of Operations Research*.

Paradise, D. B., & Courtney, J. F. (1989). Organizational Knowledge Management. *Information Resources Management Journal*, *2*(3), 1-13.

Pracht, W. E. (1990). Model visualization: Graphical support for DSS problem structuring and knowledge organization. *Decision Support Systems*, *6*(1), 13-27.

Pracht, W. E., & Courtney, J. F. (1988). The effects of an interactive graphics-based DSS to support problem structuring. *Decision Sciences, 19*(3), 598-621.

Ranganathan, C., & Sethi, V. (2000). *External IT Environment: Dimensionality and Measurement.* Paper presented at the Proceedings of the twenty first international conference on Information systems, Brisbane, Queensland, Australia.

Ritchey, T. (1998). *Fritz Zwicky, 'Morphologie' and Policy Analysis*. Paper presented at the 16th Euro Conference on Operational Analysis.

Romano, N. C., Donovan, C., Chen, H., & Nunamaker, J. F. (2003). A Methodology for Analyzing Web-based Qualitative Data. *Journal of Management Information Systems*, *19*(4), 213-246.

Rondeau, P., & Litteral, L. (2001). Evolution of Manufacturing Planning and Control Systems: from Reorder Point to Enterprise Resource Planning. *Production and Inventory Management Journal, 40*(2), 1-7.

Schmidt, E. (Writer) (2006). Google: Model of a Modern Company, *Live broadcast from Palo Alto, CA. Online participation across global virtual panel* Google Video.

Seiford, L. M. (1996). Data envelopment analysis: The evolution of the state of the art (1978–1995). *Journal of Productivity Analysis, 7*(2-3), 99-137.

Simon, H. (1960). *The New Science of Management Decision*. New York: Englewood Cliffs, N.J. : Prentice-Hall, c1977.

Woods, D. D., & Patterson, E. S. (1998). *Can we ever escape from data overload? A cognitive systems diagnosis*: Cognitive systems engineering laboratory. The Ohio State University.

Wudka, J. (1998). What is Ockham's Razor. Retrieved March, 2007, from http://physics.ucr.edu/~wudka/Physics7/Notes_www/node10.html

Yager, R., Kacprzyk, J., & Fedrizzi, M. (Eds.). (1994). *Advances in the Dempster-Shafer theory of evidence*. New York: Wiley.

Zhao, M.-Y. (2006). Multiple criteria data envelopment analysis for full ranking units associated to environment impact assessment. *International Journal of Environment and Pollution*, *28*(n 3-4), 448-464.

Zwicky, F. F. (1948). Morphological Astronomy. *The Observatory, 68*(845), 121-143.

Zwicky, F. F. (1969). *Discovery, invention, research through the morphological approach*. New York: Macmillan.